

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 21:54:41 ; Search time 9294 Seconds
(without alignment)

11215.838 Million cell updates/sec

Title: US-10-030-226-1

Sequence: 1 gcactcctccgggttctg.....tccagctgtctctgacctac 2405

Scoring table: IDENTITY_NUC

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

```

Maximum DB seq length: 20000000000

```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1:	gb_ba.*
2:	gb_hng.*
3:	gb_in.*
4:	gb_om.*
5:	gb_ov.*
6:	gb_pac.*
7:	gb_ph.*
8:	gb_pl.*
9:	gb_pr.*
10:	gb_ro.*
11:	gb_sts.*
12:	gb_sy.*
13:	gb_un.*
14:	gb_va.*
15:	em_ba.*
16:	em_fun.*
17:	em_hum.*
18:	em_in.*
19:	em_mu.*
20:	em_om.*
21:	em_ov.*
22:	em_ov.*
23:	em_pat.*
24:	em_pl.*
25:	em_pl.*
26:	em_ro.*
27:	em_sts.*
28:	em_un.*
29:	em_vi.*
30:	em_hng_hum.*
31:	em_hng_inv.*
32:	em_hng_other.*
33:	em_hng_mus.*
34:	em_hng_pun.*
35:	em_hng_rod.*
36:	em_hng_mam.*
37:	em_hng_vrt.*
38:	em_sy.*
39:	em_hng_hum.*
40:	em_hng_mus.*
41:	em_hng_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result		Query			DB		ID	Description
No.	Score	Match	Length	DB				
1	2405	100.0	2405	6	BD012071	BD012071	Fatty aci	
2	2405	100.0	2405	9	AK075377	AK075377	Homo sapi	
3	2405	100.0	2430	6	AX677203	AX677203	Sequence	
4	2405	100.0	2430	6	BC003041	BC003041	Homo sapi	
5	2405	100.0	2430	6	BC009916	BC009916	Homo sapi	
6	2392	99.5	2402	6	AX101424	AX101424	Sequence	
7	2387	99.3	2414	6	AY358409	AY358409	Homo sapi	
8	2387	99.3	2574	6	AX573775	AX573775	Sequence	
9	2387	99.3	2343	9	AK125102	AK125102	Homo sapi	
10	2310.8	96.1	2314	6	AX882240	AX882240	Sequence	
11	2310.8	96.1	2314	6	BD159613	BD159613	Primer fo	
12	2310.8	96.1	2314	9	AK027992	AK027992	Homo sapi	
13	2280	94.8	2334	9	BC029759	BC029759	Homo sapi	
14	2119.4	88.1	2166	6	AR168816	AR168816	Sequence	
15	2119.4	88.1	2166	6	AR193703	AR193703	Sequence	
16	2119.4	88.1	2166	6	AR369627	AR369627	Sequence	
17	2119.4	88.1	2166	6	AX101409	AX101409	Sequence	
18	1631.2	67.8	1833	9	AK123997	AK123997	Homo sapi	
19	1541.4	64.1	1699	6	AR338663	AR338663	Sequence	
20	1310.2	54.5	2087	6	AR168876	AR168876	Sequence	
21	1310.2	54.5	2087	6	AR193663	AR193663	Sequence	
22	1310.2	54.5	2087	6	AR369687	AR369687	Sequence	
23	1310.2	54.5	2087	6	AX101314	AX101314	Sequence	
24	1303.4	54.2	1845	10	AF077758	AF077758	Mus muscu	
25	1298.4	54.0	1998	6	AR168904	AR168904	Sequence	
26	1298.4	54.0	1998	6	AR193691	AR193691	Sequence	
27	1298.4	54.0	1998	6	AR369615	AR369615	Sequence	
28	1298.4	54.0	1998	6	AX101376	AX101376	Sequence	
29	1239.4	51.5	1380	9	BC003654	BC003654	Homo sapi	
30	1189.8	49.5	2505	10	BC060052	BC060052	Mus muscu	
31	1143	47.5	1173	6	AR168895	AR168895	Sequence	
32	1143	47.5	1173	6	AR193682	AR193682	Sequence	
33	1143	47.5	1173	6	AR369606	AR369606	Sequence	
34	1143	47.5	1173	6	AX101358	AX101358	Sequence	
35	866	35.0	150166	9	AL513523	AL513523	Human DNA	
36	853	35.5	4507	9	AK074134	AK074134	Homo sapi	
37	823.6	34.2	958	6	AX364559	AX364559	Sequence	
38	683.6	28.4	753	6	AR168880	AR168880	Sequence	
39	683.6	28.4	753	6	AR193667	AR193667	Sequence	
40	683.6	28.4	753	6	AR369591	AR369591	Sequence	
41	683.6	28.4	753	6	AX101332	AX101332	Sequence	
42	638	26.5	665	6	AX867499	AX867499	Sequence	
43	638	26.5	665	6	BD147641	BD147641	Primer fo	
44	453.6	18.9	2347	9	AF064355	AF064355	Homo sapi	
45	453.6	18.9	2361	9	AK123036	AK123036	Homo sapi	

ALIGNMENTS

RESULT 1	BD012071	BD012071	BD012071	2405 bp	DNA	linear	PAT 02-AUG-2002
LOCUS	BD012071						
DEFINITION	Fatty acid transfer protein, and a gene encoding the protein.						
ACCESSION	BD012071						
VERSION	BD012071.1	GI:22093260					
KEYWORDS	WO 0104301-A/1.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
EUCAST	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;						
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
AUTHORS	1 (bases 1 to 2405)						
	Morikawa,N., Masuko,Y., Ota,T., Isogai,T., Nishikawa,T. and						
	Kawai,Y.						
TITLE	Fatty acid transfer protein, and a gene encoding the protein						

JOURNAL
 Patent: WO 0104301-A 1 18-JAN-2001;
 HELIX RESEARCH INSTITUTE, NORIYUKI MORIKAWA, YASUHIKO MASUHO, OSHIO
 OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, YURI KAWAI
 OS Homo sapiens (human)
 PN WO 0104301-A/1

COMMENT
 PD 18-JAN-2001
 PR 07-JUL-2000 WO 2000JP004549
 PR 08-JUL-1999 JP 99P 194179, 25-APR-2000 JP 00P 128993 PR
 18-OCT-1999 US 60/159586
 PI NORIYUKI MORIKAWA, YASUHIKO MASUHO, TOSHIO OTA, TAKAO ISOGAI PI
 TETSUO NISHIKAWA,
 PI YURI KAWAI
 PC C12N15/12, C12N5/10, C12N1/15, C12N1/19, C12N1/21, C12P21/02, PC
 C07K14/47
 PC C07K16/18, C12Q1/02, C12Q1/68

FEATURES
 source
 FT Key Location/Qualifiers
 CDS 1..2405
 Location/Qualifiers
 (59)..(2248).
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2405; DB 6; Length 2405;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCACTCTCCCGGCTTTTCTGCTCTCCGCGCGTGTGAGTGTGAGGCGCTTGGGTAAT 60
DB 1 GCACCTCTCCGCGGCTTTTCTGCTCTCCGCGCGTGTGAGTGTGAGGCGCTTGGGTAAT 60
QY 61 GGGGCTGTGCGAGCGCGCGCGCGCTCCCTGGAAGAGATCTAGCTAGAACGAGGCGC 120
DB 61 GGGGCTGTGCGAGCGCGCGCGCGCTCCCTGGAAGAGATCTAGCTAGAACGAGGCGC 120
QY 121 CCTAGGTTTTCGAGAGGAGAGATCAAGGATGTTTTCGAGCGGCTGGAACGAGCGTCC 180
DB 121 CCTAGGTTTTCGAGAGGAGAGATCAAGGATGTTTTCGAGCGGCTGGAACGAGCGTCC 180
QY 181 GATGAGGAGAGCGGCTCCAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
DB 181 GATGAGGAGAGCGGCTCCAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 241 GCTGCTGTGCTGAGCTACACCTCTGCGCGCACTGCGCTGCGCTTCCGCGGACTTGGC 300
DB 241 GCTGCTGTGCTGAGCTACACCTCTGCGCGCACTGCGCTGCGCTTCCGCGGACTTGGC 300
QY 301 CTTTGCGGTGAGACTCTGTGCTGCAAAAGGCTCTTTCGAGCTGCGCGCTTGGCGGCG 360
DB 301 CTTTGCGGTGAGACTCTGTGCTGCAAAAGGCTCTTTCGAGCTGCGCGCTTGGCGGCG 360
QY 361 TGCGCGGAGCGCGGAGAGATCCCGAGGGGCTGAGCGCTGAGCGCTGAGCGCTGAGACT 420
DB 361 TGCGCGGAGCGCGGAGAGATCCCGAGGGGCTGAGCGCTGAGCGCTGAGCGCTGAGACT 420
QY 421 GGGCCAGAGCGCGCGCGCGCAACCTTTCTCATTTCAAGCTGCGGGGCTTTAGCTACTC 480
DB 421 GGGCCAGAGCGCGCGCGCGCAACCTTTCTCATTTCAAGCTGCGGGGCTTTAGCTACTC 480
QY 481 AGAGCGGAGCGCGGAGAGTAAAGGAGCTGACACCGGCTTCTCTAGTGGCTAGGCTGGGA 540
DB 481 AGAGCGGAGCGCGGAGAGTAAAGGAGCTGACACCGGCTTCTCTAGTGGCTAGGCTGGGA 540
QY 541 CTGGGAGACCGGACCGGCGCGCAACGCGGCGAGGGGAGCGCTGAGGAGGCGAGCGGCGAG 600
DB 541 CTGGGAGACCGGACCGGCGCGCAACGCGGCGAGGGGAGCGCTGAGGAGGCGAGCGGCGAG 600
QY 601 GCGGAGAGCGGAGATCAAGCGCGCGGAGCGCGGAGCTTTTCGCGAGGAGGAGCGCTGC 660
DB 601 GCGGAGAGCGGAGATCAAGCGCGCGGAGCGCGGAGCTTTTCGCGAGGAGGAGCGCTGC 660

```

```

QY 661 CGCCAGAGGTGAGAGGAGCGCGCGCGCTCTGTCACTGTGAGCACTGTGCGCTCTCT 720
DB 661 CGCCAGAGGTGAGAGGAGCGCGCGCGCTCTGTCACTGTGAGCACTGTGCGCTCTCTCT 720
QY 721 CCGCGCTGCGCGAGAGTTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 780
DB 721 CCGCGCTGCGCGAGAGTTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 780
QY 781 TGCGTTTGTGCGGACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
DB 781 TGCGTTTGTGCGGACCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
QY 841 CGCGCGCGCGCGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 900
DB 841 CGCGCGCGCGCGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 900
QY 901 CCGGAGCGCGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 960
DB 901 CCGGAGCGCGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 960
QY 961 TAGCGATTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1020
DB 961 TAGCGATTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1020
QY 1021 TTCCCGCCAGAGCATTAACAGACAGCGTGCCTGTACATCTTCACTTGTGCAACAGG 1080
DB 1021 TTCCCGCCAGAGCATTAACAGACAGCGTGCCTGTACATCTTCACTTGTGCAACAGG 1080
QY 1081 CCCCAAGGCTGTGCGATGATCATCTGAAGATCTGCAATGCGAGGCTTTCTATAGCT 1140
DB 1081 CCCCAAGGCTGTGCGATGATCATCTGAAGATCTGCAATGCGAGGCTTTCTATAGCT 1140
QY 1141 GTGTGTGTGTCACAGAGAGATGTATCTACTGCGCTCCCACTTCACTTCACTTCACT 1200
DB 1141 GTGTGTGTGTCACAGAGAGATGTATCTACTGCGCTCCCACTTCACTTCACTTCACT 1200
QY 1201 TTCCCTGTGCGCATCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCG 1260
DB 1201 TTCCCTGTGCGCATCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCG 1260
QY 1261 GTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1320
DB 1261 GTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1320
QY 1321 CATTTGGGAGAGCTGTGCTGCGATCTTGTCAACAGCGCGCGCGAGAGAGAGAGAG 1380
DB 1321 CATTTGGGAGAGCTGTGCTGCGATCTTGTCAACAGCGCGCGCGAGAGAGAGAGAG 1380
QY 1381 TAAAGTCCGCGCTGTGCGAGTGGGCAAGCGGCTGCGCGCAATCTTGTGCGG 1440
DB 1381 TAAAGTCCGCGCTGTGCGAGTGGGCAAGCGGCTGCGCGCAATCTTGTGCGG 1440
QY 1441 GCGCTTGGGCGCGCTGCGAGTGGGCAAGCGGCTGCGCGCAATCTTGTGCGG 1500
DB 1441 GCGCTTGGGCGCGCTGCGAGTGGGCAAGCGGCTGCGCGCAATCTTGTGCGG 1500
QY 1501 CATCAACTTCAACAGGAGAGCGGCGCTGTGCGGCTGTGCGGCTTCAAGCATAT 1560
DB 1501 CATCAACTTCAACAGGAGAGCGGCGCTGTGCGGCTGTGCGGCTTCAAGCATAT 1560
QY 1561 CTTCCCTTCTCTGTGATGCTATGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1561 CTTCCCTTCTCTGTGATGCTATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 GGGGCACTGTATGCGCATCTTCAAGGAGAGCGGCTGTGCGGCGCGCGGATGAGCA 1680
DB 1621 GGGGCACTGTATGCGCATCTTCAAGGAGAGCGGCTGTGCGGCGCGCGGATGAGCA 1680
QY 1681 GCAATGCCCATCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1740
DB 1681 GCAATGCCCATCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1740
QY 1741 GGAATGCTTTCGCGCTGCGGAGATGTTTCTTCAACACTGCGGAGCTGTGCTGAGTGA 1800

```

Db 1741 GGATGCTTCGCGGCTGGGAGATGTTTCTTCAACACCTGGGAGCTGTGTCTGCCATGA 1800
Qy 1801 CCAAGGTTTTCTCGGCTTCCATGATCGTACTGGAGACACCTTCAAGTGGAGGGGAGAA 1860
Db 1801 CCAAGGTTTTCTCGGCTTCCATGATCGTACTGGAGACACCTTCAAGTGGAGGGGAGAA 1860
Qy 1861 TGTGGCCACACCGAGGTGGCAGAGGTCTTCCAGGCCCCCTGATTTTCTTACAGAGTGA 1920
Db 1861 TGTGGCCACACCGAGGTGGCAGAGGTCTTCCAGGCCCCCTGATTTTCTTACAGAGTGA 1920
Qy 1921 CGTGTATGAGTCACTGTGCGCAGGGCATGAAGGAGGCTGATGATGAGCCCTAGTTCT 1980
Db 1921 CGTGTATGAGTCACTGTGCGCAGGGCATGAAGGAGGCTGATGATGAGCCCTAGTTCT 1980
Qy 1981 GCGTCCCGCCCAAGCTTTGAGACCTTATGACAGCTTACACCCAGTGTCTGAGAACTTGGC 2040
Db 1981 GCGTCCCGCCCAAGCTTTGAGACCTTATGACAGCTTACACCCAGTGTCTGAGAACTTGGC 2040
Qy 2041 ACCTTATGCGCGGCGCGGATTCCTCAGAGCTCCAGAGCTCTTTGGCCACACAGACCTT 2100
Db 2041 ACCTTATGCGCGGCGCGGATTCCTCAGAGCTCCAGAGCTCTTTGGCCACACAGACCTT 2100
Qy 2101 CAAACAGCAGAAAGTTGCGATGAGCAATGAGGGCTTGACCCGACACCTGTCTGACCC 2160
Db 2101 CAAACAGCAGAAAGTTGCGATGAGCAATGAGGGCTTGACCCGACACCTGTCTGACCC 2160
Qy 2161 ACTGTACGTTCTGACAGAGGTGATGAGTCTACCTGCCCCCTCAACCTGCGGATACAG 2220
Db 2161 ACTGTACGTTCTGACAGAGGTGATGAGTCTACCTGCCCCCTCAACCTGCGGATACAG 2220
Qy 2221 CGCCCTCTGCGCAGAAACCTTCGAATCTGAGAACCTTCAACCTGAGGACCTGAGAGA 2280
Db 2221 CGCCCTCTGCGCAGAAACCTTCGAATCTGAGAACCTTCAACCTGAGGACCTGAGAGA 2280
Qy 2281 GGAAGCTCTGAGGGGTGGGGGCGCTTGCAGAGTGTACTGAGGCTGTGAGGAGATCTTTCTATA 2340
Db 2281 GGAAGCTCTGAGGGGTGGGGGCGCTTGCAGAGTGTACTGAGGCTGTGAGGAGATCTTTCTATA 2340
Qy 2341 CCAAGACTGCGGCTCACTATTGTAATTAATAGTGGCTGAGAGCTGATCAGCTGTCTGTGA 2400
Db 2341 CCAAGACTGCGGCTCACTATTGTAATTAATAGTGGCTGAGAGCTGATCAGCTGTCTGTGA 2400
Qy 2401 CCTAC 2405
Db 2401 CCTAC 2405

RESULT 2
AK075377 2405 bp mRNA linear PRI 03-SEP-2002
LOCUS Homo sapiens cDNA PSEC067 f1s, clone NT2RP2001142, weakly similar
DEFINITION to VERY-LONG-CHAIN ACYL-COA SYNTHETASE (EC 6.2.1.-).
ACCESSION AK075377
VERSION AK075377.1 GI:22761422
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Nishikawa, T., Suzuki, Y., Kawai-Hio, Y., Hayaishi, K.,
Ishii, S., Saito, K., Yamamoto, J., Nakamura, A., Nagai, T.,
Nakamura, Y., Nagahara, K., Sugano, S. and Isogai, T.
TITLE HRI human cDNA sequencing project
JOURNAL Unpublished
AUTHORS Isogai, T. and Yamamoto, J.
TITLE 2 (bases 1 to 2405)
JOURNAL Direct Submission
AUTHORS Submitted (20-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 229-0812, Japan
COMMENT (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
HRI human cDNA sequencing project; cDNA 5'- & 3'-end one pass

sequencing, clone selection and full insert sequencing: Helix
Research Institute (supported by Japan Key Technology Center etc.);
CDNA library construction: Institute of Medical Science, University
of Tokyo, Laboratory of Genome Structure, Human Genome Center.
location/Qualifiers
1. 2405
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2001142"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_1b="NT2R2"
/note="Cloning vector: pME18SFu3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction"
59..2251
/note="unamed protein product"
/codon_start=1
/protein_id="BAC11578.1"
/db_xref="GI:22761423"
/translation="MGVCGRTAPKPKKESQLEALGFRKGSQWFAQGNQVPIE
AGSMALILPLILPLILKILHLPOLRMLPLDARVAVLCCRRALRALAA
ADPEPBECCSLARLALALAGRAATFLHGRFRFSSEAPRSNRAAPLPLALG
DMGPDGSDSGESAGEERAAAGADDAAGSAGEFAGDGAARGGALAPLSPGATV
LLPAPPEFLMLPGLARARATVPTALRVRGPHLHLRSCGALVALVAPESLLE
PDLPLRAMGLHMLVAGTGPAGISDLIAVSEVDPVPGVLSPPSISIDTCLYIF
TSQTLGPRAAISHKILICCGFYOLGDEVDIYALPLYNHSGSLIIVGCMGI
GATVILKSPAGSGFEDCOOHRTVTFVYIISLCGYLVNOPSXAEKRGKXVLA
LRPDTWESFVRREGPIQYLETGTEGNVATIGTGAGVAGRAVSWLXKHPPEFLIR
IDVTGTEPRDQHCMAISREBRLVLAIPISQSPFLGVAGSEPLAQLKQVFR
GVFENPDILVCDQDGLRPHRDRTQFLRMKSGEVATVYAEVPEADDTLOEVAVIC
VTPGHEBRAGMAALVLRPPALDLMQYTVHSENLPPYARPRFLRDSLATVETFK
QOKVMAWNEGFPSTLSDPLVYLDAVAVLPLTARYSALLAGNLRI"

ORIGIN
Query Match 100.0%; Score 2405; DB 9; Length 2405;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
Matches 2405; Conservative 0; Indels 0; Gaps 0;
Qy 1 GCACTCTCCCGGGTTTCTGCTCTCGCCCTGTGAGTGTGGGGCTGTGGTGGGAAT 60
Db 1 GCACTCTCCCGGGTTTCTGCTCTCGCCCTGTGAGTGTGGGGCTGTGGTGGGAAT 60
Qy 61 GGGCTGTGCCACCGCAGCCGGCTCCCTGGAAGAGAAAGTCTACGTAGAACGAGCGGC 120
Db 61 GGGCTGTGCCACCGCAGCCGGCTCCCTGGAAGAGAAAGTCTACGTAGAACGAGCGGC 120
Qy 121 CCTAGTTTGGAGAGGAGATCAGAGGATGTTTGGAGCGGCTGGAACGAGGTGCC 180
Db 121 CCTAGTTTGGAGAGGAGATCAGAGGATGTTTGGAGCGGCTGGAACGAGGTGCC 180
Qy 181 GATAGAGAGAGCGGGCTCATGAGTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTAC 240
Db 181 GATAGAGAGAGCGGGCTCATGAGTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTAC 240
Qy 241 GCTGCTGCTGCTGAGACTACACCTCTGAGCGAGTGTGAGCTGCGGCTGCGGAGCTTGGC 300
Db 241 GCTGCTGCTGCTGAGACTACACCTCTGAGCGAGTGTGAGCTGCGGCTGCGGAGCTTGGC 300
Qy 301 CTTTGGGTCGAGCTCTGTGCTGCAAAAAGGCTCTTTCAGCTCGGCGCTTGCAGCGCGGC 360
Db 301 CTTTGGGTCGAGCTCTGTGCTGCAAAAAGGCTCTTTCAGCTCGGCGCTTGCAGCGCGGC 360
Qy 361 TGTCCGCCAACCGGAAAGTTCGAGAGGGGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 420
Db 361 TGTCCGCCAACCGGAAAGTTCGAGAGGGGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 420
Qy 421 GGGCCAGAGCGGCGCGGCACACCTTCTCATTCACGAGCTGCGGCGCTTTCAGCTACTC 480
Db 421 GGGCCAGAGCGGCGCGGCACACCTTCTCATTCACGAGCTGCGGCGCTTTCAGCTACTC 480
Qy 481 AGAGGCGAGCGCGAGAGTACAGGAGCTGACGCGCTTCTCACTGAGCTGAGCTGAGG 540

Db	1561	CTTCCCCCTTCCTTGATGTGCTATGATGATGATCACCACAGAGAGGCCAATTTGGGACC	1620
Qy	1621	GGGGCACACTGATGAGGCCAATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCA	1680
Db	1621	GGGGCACACTGATGAGGCCAATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCA	1680
Qy	1681	GCACTCCCATCTCTGGGCGATAGCTGAGCGGGCCAGAGCTGGCCCCAGGGGAAATTTGCTAAA	1740
Db	1681	GCACTCCCATCTCTGGGCGATAGCTGAGCGGGCCAGAGCTGGCCCCAGGGGAAATTTGCTAAA	1740
Qy	1741	GGATGTCCTCCGGCTGGGGATGTTTCTTCAACACTGGGGGACCTGCTGGTCTGCGATGA	1800
Db	1741	GGATGTCCTCCGGCTGGGGATGTTTCTTCAACACTGGGGGACCTGCTGGTCTGCGATGA	1800
Qy	1801	CCAAAGTTTTCTCCGCTTCCATGATGCTACTGAGACACCTTCAGTGGAAAGGGGGAA	1860
Db	1801	CCAAAGTTTTCTCCGCTTCCATGATGCTACTGAGACACCTTCAGTGGAAAGGGGGAA	1860
Qy	1861	TGTGGCCACAACCGAGGTGGGACAGGCTTCGAGGCCCTTGATTTTCTTACAGGAGTGAA	1920
Db	1861	TGTGGCCACAACCGAGGTGGGACAGGCTTCGAGGCCCTTGATTTTCTTACAGGAGTGAA	1920
Qy	1921	CGTCTATGAGATCATCTGTGCCAAGGACATGAAGCAGGGCTGGAATGGCAGCCCTTATGTTCT	1980
Db	1921	CGTCTATGAGATCATCTGTGCCAAGGACATGAAGCAGGGCTGGAATGGCAGCCCTTATGTTCT	1980
Qy	1981	GGGTCCCCCCCCCAGCTTTGGACCTTATGACAGCTCTACCCAGCTGTCTAGAACTTGGC	2040
Db	1981	GGGTCCCCCCCCCAGCTTTGGACCTTATGACAGCTCTACCCAGCTGTCTAGAACTTGGC	2040
Qy	2041	ACCTTATGCCCCCGGCCGATTCTCTCAGAGCTCCAGAGCTCTTTGGCCACCAAGAGACCTT	2100
Db	2041	ACCTTATGCCCCCGGCCGATTCTCTCAGAGCTCCAGAGCTCTTTGGCCACCAAGAGACCTT	2100
Qy	2101	CAAAACAGCAGAAAGTTGCGATGACAAATGAGGGCTTCGACCCAGCACCCTGTCTGACC	2160
Db	2101	CAAAACAGCAGAAAGTTGCGATGACAAATGAGGGCTTCGACCCAGCACCCTGTCTGACC	2160
Qy	2161	ACTGTACGTTCTGACCAAGGCTGTAGGTGCTCTACCTGGCCCTTCAACACTGCCCGGTACAG	2220
Db	2161	ACTGTACGTTCTGACCAAGGCTGTAGGTGCTCTACCTGGCCCTTCAACACTGCCCGGTACAG	2220
Qy	2221	CGCCCTCTCCGGCAGAGAAACCTTGAATCTGAGAACTTCCACACTAGAGGACCTGAGAGA	2280
Db	2221	CGCCCTCTCCGGCAGAGAAACCTTGAATCTGAGAACTTCCACACTAGAGGACCTGAGAGA	2280
Qy	2281	GGAACCTCTGTGGGGGTGGGGCCGTGACAGTGTACTGGGCTGTCAAGGATCTTTTCTATA	2340
Db	2281	GGAACCTCTGTGGGGGTGGGGCCGTGACAGTGTACTGGGCTGTCAAGGATCTTTTCTATA	2340
Qy	2341	CCAGAACTGGGCTCACTATTTTGTATAATGAGCTGGAGCTGATCCAGCTGTCTCTGA	2400
Db	2341	CCAGAACTGGGCTCACTATTTTGTATAATGAGCTGGAGCTGATCCAGCTGTCTCTGA	2400
Qy	2401	CTTAC 2405	
Db	2401	CTTAC 2405	

RESULT	3
LOCUS	AX677203
DEFINITION	AX677203 2430 bp DNA linear PAT 27-MAR-2003
ACCESSION	Sequence 396 from Patent WO02103028.
VERSION	AX677203
KEYWORDS	AX677203.1 GI:29334629
SOURCE	.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 Barenova, A.V., Yankovsky, N.K., Kozlov, A.P., Lobashev, A.V. and Krukshaya, L.L.

TITLE In silico screening for phenotype-associated expressed sequences
JOURNAL Patent: WO 02103028-A 396 27-DEC-2002;
Biomedical Center (RU)

FEATURES Location/Qualifiers

Source
1. 2430
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2405; DB 6; Length 2430;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCACCTCCCTCCGGGTTTCTGCTCTCCGCCGTGTGAGTGTGGGGCTTGGTGGGAT 60
DB 8 GCACCTCCCTCCGGGTTTCTGCTCTCCGCCGTGTGAGTGTGGGGCTTGGTGGGAT 67
QY 61 GGGGCTGTGCCAGCGCAAGCGCGCTCCCTGGAGAGAAATCTCAGCTAGAGAGCGGC 120
DB 68 GGGGCTGTGCCAGCGCAAGCGCGCTCCCTGGAGAGAAATCTCAGCTAGAGAGCGGC 127
QY 121 CCTAGGTTTGGAGAGGAGATCAGGAGTGTTCGAGCGGCTGAAACGACAGCGTGCC 180
DB 128 CTTAGGTTTGGAGAGGAGATCAGGAGTGTTCGAGCGGCTGAAACGACAGCGTGCC 187
QY 181 GATAGAGAGAGCGGCTTCATAGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 188 GATAGAGAGAGCGGCTTCATAGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 247
QY 241 GCTGCTGCTGCTGAGAGCTACACTCTGAGCGCAGTTGAGCGCTTCCGGCGACTGGC 300
DB 248 GCTGCTGCTGCTGAGAGCTACACTCTGAGCGCAGTTGAGCGCTTCCGGCGACTGGC 307
QY 301 CTTTGCAGTGCAGCTCTGTGCTGTCAAAAGGACTCTTTCAGACTGCGCCCTGCGCGC 360
DB 308 CTTTGCAGTGCAGCTCTGTGCTGTCAAAAGGACTCTTTCAGACTGCGCCCTGCGCGC 367
QY 361 TGCCTGCGCAATCCCGAAGGTCCCGAGAGGGGCTGAGAGCTGAGCGGCTGCGGAGCT 420
DB 368 TGCCTGCGCAATCCCGAAGGTCCCGAGAGGGGCTGAGAGCTGAGCGGCTGCGGAGCT 427
QY 421 GGCCTCAGAGCGCGCGCGCACACCTTCTCATTTACAGCGCTGCGCGCTTTTACCTATC 480
DB 428 GGCCTCAGAGCGCGCGCGCACACCTTCTCATTTACAGCGCTGCGCGCTTTTACCTATC 487
QY 481 AGAGCGGAGCGCGAGAGTAAACAGGAGCTGCAAGCGGCTTCTTAAGTGTGCTAAGCTGGA 540
DB 488 AGAGCGGAGCGCGAGAGTAAACAGGAGCTGCAAGCGGCTTCTTAAGTGTGCTAAGCTGGA 547
QY 541 CTGGGAGACCCGACGCGCGGACAGCGCGAGGAGCGCTGAGAGAGCGAGCGGAGC 600
DB 548 CTGGGAGACCCGACGCGCGGACAGCGCGAGGAGCGCTGAGAGAGCGAGCGGAGC 607
QY 601 GCGCGAGACCCGAGATGACAGCGCGGAGAGCGCGGAGTTTGTCCGAGAGGAGAGCTGTC 660
DB 608 GCGCGAGACCCGAGATGACAGCGCGGAGAGCGCGGAGTTTGTCCGAGAGGAGAGCTGTC 667
QY 661 CGCCAGAGTGGAGAGCGCGCGCTCTGTGCTCAGCTGAGAGAAATGTGAGCGCTGCTCT 720
DB 668 CGCCAGAGTGGAGAGCGCGCGCTCTGTGCTCAGCTGAGAGAAATGTGAGCGCTGCTCT 727
QY 721 CCCCGCTGGCCAGAGTTTCTGTGCTCTGTGCTTGGGCTTGGCCAGAGCGCGCTTGGCAG 780
DB 728 CCCCGCTGGCCAGAGTTTCTGTGCTCTGTGCTTGGGCTTGGCCAGAGCGCGCTTGGCAG 787
QY 781 TGCCTTTTGCCCAATCCCGCTGCGCGCGGCGCGCTGCTGCACTGCGAGCTGAGG 840
DB 788 TGCCTTTTGCCCAATCCCGCTGCGCGGCGCGCTGCTGCACTGCGAGCTGAGG 847
QY 841 CGCGCGCGCGCTGAGTGTGCGCGCAGAGTTTCTGAGAGTCCCTGAGCGCGAGCTTCCCGC 900
DB 848 CGCGCGCGCGCTGAGTGTGCGCGCAGAGTTTCTGAGAGTCCCTGAGCGCGAGCTTCCCGC 907

```

```

QY 901 CCTGAGAGCCATGGGGCTTCACCTTGTGGCTGACAGGCCAGAGAACCCCTGCTGAGAT 960
DB 908 CTTGAGAGCCATGGGGCTTCACCTTGTGGCTGACAGGCCAGAGAACCCCTGCTGAGAT 967
QY 961 TAGGATTTTGGGCTGAGAGTGTCCGCTGAGTGTGAGGCGAGTGCAGATACCTCTC 1020
DB 968 TAGGATTTTGGGCTGAGAGTGTCCGCTGAGTGTGAGGCGAGTGCAGATACCTCTC 1027
QY 1021 TTCCCCCAGAGCATPAACAGACGTCCTGTATCATCTTCACTTGTGGCACACCGGCT 1080
DB 1028 TTCCCCCAGAGCATPAACAGACGTCCTGTATCATCTTCACTTGTGGCACACCGGCT 1087
QY 1081 CCCCAAGGCTGTGGATTCAGTATCTGAAAGTCTGTGAATGCCAGGCTTATACAGCT 1140
DB 1088 CCCCAAGGCTGTGGATTCAGTATCTGAAAGTCTGTGAATGCCAGGCTTATACAGCT 1147
QY 1141 GTGTGTGTCCACAGAGAGATGTATCTACCTGCGCTCCACTCTACACATATCCG 1200
DB 1148 GTGTGTGTCCACAGAGAGATGTATCTACCTGCGCTCCACTCTACACATATCCG 1207
QY 1201 TTCCCTGCTGGCATGTGTGGCTGATGGGCAATGGGGCAAGTGTGTAATCCAA 1260
DB 1208 TTCCCTGCTGGCATGTGTGGCTGATGGGCAATGGGGCAAGTGTGTAATCCAA 1267
QY 1261 GTTCTGCGCTGATCTGAGGAGATTCTGAGGAGATTGACAGAGACAGGTTGTCAGTA 1320
DB 1268 GTTCTGCGCTGATCTGAGGAGATTCTGAGGAGATTGACAGAGACAGGTTGTCAGTA 1327
QY 1321 CATTTGGAGAGCTGTGCGGATCTTGTCAACAGCCCGGAGAGAGCGAGAACTGTGGCA 1380
DB 1328 CATTTGGAGAGCTGTGCGGATCTTGTCAACAGCCCGGAGAGAGCGAGAACTGTGGCA 1387
QY 1381 TAAGGTCCGCTGGAGAGTGGGAGCGGCGTGGCGCCAGATCTGAGGAGCTTTTGGG 1440
DB 1388 TAAGGTCCGCTGGAGAGTGGGAGCGGCGTGGCGCCAGATCTGAGGAGCTTTTGGG 1447
QY 1441 GCGCTTCGCGCCCTGACAGTGTCTGAGACATATGAGTCTGACAGAGGCACTGTGGCAC 1500
DB 1448 GCGCTTCGCGCCCTGACAGTGTCTGAGACATATGAGTCTGACAGAGGCACTGTGGCAC 1507
QY 1501 CATCACTACACAGAGCAACCGGGCGCTGTGGGGGTGCTTCTGCTTACAGAGATAT 1560
DB 1508 CATCACTACACAGAGCAACCGGGCGCTGTGGGGGTGCTTCTGCTTACAGAGATAT 1567
QY 1561 CTTCCCTTCTCCTTGAATTCGATATGATGACACAGAGAGCCAAATCGGAGCCCGCA 1620
DB 1568 CTTCCCTTCTCCTTGAATTCGATATGATGACACAGAGAGCCAAATCGGAGCCCGCA 1627
QY 1621 GGGGCACTGTATGAGCCACATCTTCAGGTAGCCAGGGCTGTGTGGCCCCGGTAAAGCA 1680
DB 1628 GGGGCACTGTATGAGCCACATCTTCAGGTAGCCAGGGCTGTGTGGCCCCGGTAAAGCA 1687
QY 1681 GCACTCCCAATCTGTGGCTATGCTGAGCGGCGCAGAGCTGCGCCAGGGAGTTGCTAAA 1740
DB 1688 GCACTCCCAATCTGTGGCTATGCTGAGCGGCGCAGAGCTGCGCCAGGGAGTTGCTAAA 1747
QY 1741 GATGTCTTCGCGCTGTGGGATTTTCTTCAACACTGGGAGCTCTGTGTGCGATGA 1800
DB 1748 GATGTCTTCGCGCTGTGGGATTTTCTTCAACACTGGGAGCTCTGTGTGCGATGA 1807
QY 1801 CCAAGGTTTCTCCGCTTCCATATGATGTAGTGAACACTTAAAGTGAAGGGAGAA 1860
DB 1808 CCAAGGTTTCTCCGCTTCCATATGATGTAGTGAACACTTAAAGTGAAGGGAGAA 1867
QY 1861 TGTGTCCACACAGAGTGTGACAGAGTCTTGAAGCCCTTGAATTTCTTCAAGAGTGA 1920
DB 1868 TGTGTCCACACAGAGTGTGACAGAGTCTTGAAGCCCTTGAATTTCTTCAAGAGTGA 1927
QY 1921 CGTCTATGAGTCACTGTGCAAGGCAATGAAGGCAAGGCTGGAATGAGCCCTAGTTCT 1980
DB 1928 CGTCTATGAGTCACTGTGCAAGGCAATGAAGGCAAGGCTGGAATGAGCCCTAGTTCT 1987

```

QY 1981 GCGTCCCGCCCAAGCTTTGAGACCTTATGACGCTCTACACCAAGTGTCTGAGAACTTGCC 2040
 Bc003041
 Db 1988 GCGTCCCGCCCAAGCTTTGAGACCTTATGACGCTCTACACCAAGTGTCTGAGAACTTGCC 2047
 QY 2041 ACTTATGCGCCGCGCCCGAATCTCTCAAGGCTTCAAGAGTCTTTGGCCACCAAGAGACCTT 2100
 Db 2048 ACTTATGCGCCGCGCCCGAATCTCTCAAGGCTTCAAGAGTCTTTGGCCACCAAGAGACCTT 2107
 QY 2101 CAAACGACGAAAGTTCGATGCGCAAAATGAGGGCTTCGACCCGACGACCCCTGTCTGACCC 2160
 Db 2108 CAAACGACGAAAGTTCGATGCGCAAAATGAGGGCTTCGACCCGACGACCCCTGTCTGACCC 2167
 QY 2161 ACTGATGCTTCTGAGACGAGCTGTAGTGTCTACCTGCCCCCTCACAACTGCCCCGTACAG 2220
 Db 2168 ACTGATGCTTCTGAGACGAGCTGTAGTGTCTACCTGCCCCCTCACAACTGCCCCGTACAG 2227
 QY 2221 CGCCCCCTGAGAGGAAACCTTGAATCTGAGAACTTCCACCTGAGGACCTGAGAGA 2280
 Db 2228 CGCCCCCTGAGAGGAAACCTTGAATCTGAGAACTTCCACCTGAGGACCTGAGAGA 2287
 QY 2281 GGAACCTGTGTGGGGTGGGGCGGTGACAGTGTACTGTGGGCTGTGAGGAGTCTTTCTATA 2340
 Db 2288 GGAACCTGTGTGGGGTGGGGCGGTGACAGTGTACTGTGGGCTGTGAGGAGTCTTTCTATA 2347
 QY 2341 CGAAGACGCGGTCTCACTATTTGTATTAATGTGCTGAGCTGATTCAGCTGTCTCTGA 2400
 Db 2348 CGAAGACGCGGTCTCACTATTTGTATTAATGTGCTGAGCTGATTCAGCTGTCTCTGA 2407
 QY 2401 CCTAC 2405
 Db 2408 CCTAC 2412

RESULT 4
 Bc003041
 LOCUS
 DEFINITION
 Homo sapiens solute carrier family 27 (fatty acid transporter)
 member 3, mRNA (cdna clone MGC:4365 IMAGE:2822875), complete cds.
 ACCESSION
 Bc003041.1 GI:12804360
 VERSION
 Bc003041.1 GI:12804360
 KEYWORDS
 MGC.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 2430)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Alekshun, S.F., Zeeberg, B., Bluetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stadleiro, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheer, T.E., Brownstein, M.J., Umedin, T.B., Toshikuni, S.,
 Carninci, P., Prange, C., Kaba, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, J., Whiting, M., Madan, A., Young, A.C., Shachrenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butler, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 MEDLINE
 2288257
 PUBLISHED
 12477932
 2 (bases 1 to 2430)
 REFERENCE
 Strausberg, R.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (06-FEB-2001) National Institutes of Health, Mammalian

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca
 Steven Jones, Jennifer Amano, Ian Bosdet, Yaron Butterfield,
 Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Matheson, Candice McLeay, Steven
 Ness, Pawan Pandon, Anna-Lisa Pridhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalhus, Michael Smith, Loraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

FEATURES
 source
 clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 9 Row: n Column: 23
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 12336578.
 Location/Qualifiers
 1..2430
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:4365 IMAGE:2822875"
 /tissue_type="Lung, small cell carcinoma"
 /clone_lib="NIH MGC-7"
 /lab_host="DH10B-R"
 /note="vector: pOTB7"
 1..2430
 /gene="SLC27A3"
 /note="synonyms: FATP3, MGC4365"
 /db_xref="locusID:11000"
 /db_xref="MIM:604193"
 66..2258
 /codon_start=1
 /product="solute carrier family 27 member 3"
 /protein_id="AAH01041.1"
 /db_xref="GI:12804361"
 /db_xref="locusID:11000"
 /translation="MGVQRTAPWKEKSLERAPALGPRKSGSGMFAAGNQTPIEE
 AGSMALILPLILPLILPLILKHLMPQLRWLPDLAAVVALCKKALRALAAAA
 ADPEEGEGSGIARLAIARLAAQRAATFLHSGSEASERSSNAARVPAALGM
 DMGPRGSGSGSAGERRAPAGDAAGSGAPFAGDGAARGGAAAPLSPEATVA
 LLPLGPRFLWMLRELAKAGRTAVPALRGLPLRLKRCGKARALVALAEFLSELE
 PDLPLRLMGLWLAAGGTAPAGISDLAVSAEIVGPPVGVYSSQSLDTLYIF
 TSGTGLPKRAGISLKLIOCGFYOLCGVQEDVILALPLVHMSGLSLGIVCMGI
 GATVLEKRFSGAGWEDECOQRTVPEYIGELRVYNQPPSAEKGKRLAVSG
 LRPDLWKEFVRFGPLQVLEFYGLEGVATINTGCGAVGASMLYKHFPSLR
 GVTGGEIRDPDCCGATSPDEGLIYAPVDSQSPLYAGVAGSELAOGKILKVRP
 GVDPRNEDLLVDCDQGTFRPHDPTGDFPRKSGSNVATTEVAEFAALDELQENYV
 VTVHEGRAGMALVLRPRAIDLMQVTVHSNLPVAPRFLRQESLATEYTRK
 QOKRMANEGDPSTLSDPLVLDQAVGAVPLTAYVSALAGNLRI"
 711..2147
 /note="Caic; Region: Acyl-CoA synthetases
 (AMP-forming)/AMP-acid ligases II [lipid metabolism /
 secondary metabolites biosynthesis, transport, and
 catabolism]"
 /db_xref="CDD:COG0318"

ORIGIN
 Query Match 100.0%; Score 2405; DB 9; Length 2430;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCATCTCTCCCGGGTTTGTGCTTCGCCCGCTGTGAGTGTGGGGCGCTGGGTGGGAAT 60

Db 8 GCACTCTCCCGGGTTTCTGCTCTCCGCCGTGTGAGTGTGGGGGCTGGGTGGGAAT 67
Qy 61 GGGCGTGTGCCAGCGCAACGGCGCTCCCTGGAAAGGAAGTCTCAGTTAAGACGCGGC 120
Db 68 GGGCGTGTGCCAGCGCAACGGCGCTCCCTGGAAAGGAAGTCTCAGTTAAGACGCGGC 127
Qy 121 CCTAGTTTTCGAAAGGAGATCAGGGATGTTTTCGAGCGGCTGGAACAGACGCGTGC 180
Db 128 CTTAGTTTTCGAAAGGAGATCAGGGATGTTTTCGAGCGGCTGGAACAGACGCGTGC 187
Qy 181 GATAGAGAAAGCGGGCTTCATGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 188 GATAGAGAAAGCGGGCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247
Qy 241 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 248 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
Qy 301 CTTTGCAGTGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 308 CTTTGCAGTGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367
Qy 361 TGCCTGCGAAGCGGAGGATCCCGAGGGGGGCTGAGAGCTGCGCTGCGGCTGCGGAACT 420
Db 368 TGCCTGCGAAGCGGAGGATCCCGAGGGGGGCTGAGAGCTGCGCTGCGGCTGCGGAACT 427
Qy 421 GGGCCAGAGCGGCGCGCGGACACCTTCTCATTCAGCGCTGCGGCGCTTATAGCTATC 480
Db 428 GGGCCAGAGCGGCGCGCGGACACCTTCTCATTCAGCGCTGCGGCGCTTATAGCTATC 487
Qy 481 AGAGGCGGAGCGCGGAGGATCAAGGGGCTGCAACGGCGCTTCTTCAAGTGCCTAGCTGGGA 540
Db 488 AGAGGCGGAGCGCGGAGGATCAAGGGGCTGCAACGGCGCTTCTTCAAGTGCCTAGCTGGGA 547
Qy 541 CTGGGAGACCGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
Db 548 CTGGGAGACCGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 607
Qy 601 GCGGAGACCGGAGGATGCAAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 608 GCGGAGACCGGAGGATGCAAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 667
Qy 661 GCGGAGGATGGAAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 668 GCGGAGGATGGAAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727
Qy 721 CCGCGCTGGCCAGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 728 CCGCGCTGGCCAGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787
Qy 781 TGCCTTTTGTCCCAACCGGCTGCGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 788 TGCCTTTTGTCCCAACCGGCTGCGCGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 847
Qy 841 GCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 848 GCGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907
Qy 901 CCTGAGAGCCATGGGGCTTCCACTGCTGGGCTGCAAGGCCAGGAACCCACCTGCTGGAAAT 960
Db 908 CCTGAGAGCCATGGGGCTTCCACTGCTGGGCTGCAAGGCCAGGAACCCACCTGCTGGAAAT 967
Qy 961 TAGCATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 968 TAGCATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027
Qy 1021 TTCCCTCCAGAGCATPAACAGACAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1028 TTCCCTCCAGAGCATPAACAGACAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087
Qy 1081 CCCCAAGGTGCTCGGATCAGTCACTGAAGTCTGCAATGCCAGGGCTTCTATCAGCT 1140

Db 1088 CCCAAGGTGCTCGGATCAGTCACTGAAGTCTGCAATGCCAGGGCTTCTATCAGCT 1147
Qy 1141 GTGTGGTTCACACAGAGATGATCTACCTCGCCCTCCACTCTACCATGTCGG 1200
Db 1148 GTGTGGTTCACACAGAGATGATCTACCTCGCCCTCCACTCTACCATGTCGG 1207
Qy 1201 TTCCCTGCTGGCATGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1208 TTCCCTGCTGGCATGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267
Qy 1261 GTTCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db 1268 GTTCTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1327
Qy 1321 CATGGGAGCTGTGCGGATCTTGTCAACCAAGCCCTCGAGCAAGCAAGCGTGCCA 1380
Db 1328 CATGGGAGCTGTGCGGATCTTGTCAACCAAGCCCTCGAGCAAGCAAGCGTGCCA 1387
Qy 1381 TAAAGTCCGCTGGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1388 TAAAGTCCGCTGGAGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1447
Qy 1441 GCGCTTCGCGCCCTGCGAGGCTGCGAGACATATGAGCTGACAGAGGCAAGTGGCCAC 1500
Db 1448 GCGCTTCGCGCCCTGCGAGGCTGCGAGACATATGAGCTGACAGAGGCAAGTGGCCAC 1507
Qy 1501 CATCAACTACACAGACAGCGGGGCGCTGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 1508 CATCAACTACACAGACAGCGGGGCGCTGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1567
Qy 1561 CTTCCCTTCTGCTGATTCGCTATGATGTCACACAGAGAGCCAAATTCGGGACCCCA 1620
Db 1568 CTTCCCTTCTGCTGATTCGCTATGATGTCACACAGAGAGCCAAATTCGGGACCCCA 1627
Qy 1621 GGGGCACTGTATGGCCATCTCCAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db 1628 GGGGCACTGTATGGCCATCTCCAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1687
Qy 1681 GCAATGCCATTCCTGGGCTATGCTGCGGGGCGAGAGCTGCGGCCAGGGAGGTTGCTAA 1740
Db 1688 GCAATGCCATTCCTGGGCTATGCTGCGGGGCGAGAGCTGCGGCCAGGGAGGTTGCTAA 1747
Qy 1741 GATATCTTCGCGCTGGGAGATGTTTCTTCAACACTGGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1748 GATATCTTCGCGCTGGGAGATGTTTCTTCAACACTGGGAGCCTGCTGCTGCTGCTGCTGCTGCTGCT 1807
Qy 1801 CCAAGTTTTTCTCGCTTCATGATGCTGGAACACTTCAAGTGAAGGGGAGAA 1860
Db 1808 CCAAGTTTTTCTCGCTTCATGATGCTGGAACACTTCAAGTGAAGGGGAGAA 1867
Qy 1861 TGTGGCACAAACGAGGTCTTCAAGGCGCTTCAAGTCTTCAAGAGGTGAA 1920
Db 1868 TGTGGCACAAACGAGGTCTTCAAGGCGCTTCAAGTCTTCAAGAGGTGAA 1927
Qy 1921 CGTCTATGAGTCACTGTGCGAGGCTATGAAGCGGCTGGAATGCGACGCTTATGCT 1980
Db 1928 CGTCTATGAGTCACTGTGCGAGGCTATGAAGCGGCTGGAATGCGACGCTTATGCT 1987
Qy 1981 GCGTCCCGCCCAAGCTTGTGAACCTTATGCACTTCAACCAAGTGTGAGAACTTGGC 2040
Db 1988 GCGTCCCGCCCAAGCTTGTGAACCTTATGCACTTCAACCAAGTGTGAGAACTTGGC 2047
Qy 2041 ACCTTATGCCCGGCGGATTCCTCAAGGCTCCAGAGTCTTTGGCCACCAAGAGACTT 2100
Db 2048 ACCTTATGCCCGGCGGATTCCTCAAGGCTCCAGAGTCTTTGGCCACCAAGAGACTT 2107
Qy 2101 CAACAGCAAGAAAGTTCGATGAGCAATGAGGCTTCAAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2160
Db 2108 CAACAGCAAGAAAGTTCGATGAGCAATGAGGCTTCAAGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2167
Qy 2161 ACTGTAGTTCGAGCAAGGCTGTAGGTGCTTACCTGCTTCAACATGCGCGGTACAG 2220
Db 2168 ACTGTAGTTCGAGCAAGGCTGTAGGTGCTTACCTGCTTCAACATGCGCGGTACAG 2227

QY 2221 CGCCCTCTGGAGAGAACTTCGAACTGAGACTTCACACCTGAGCACTGAGAGA 2280
 DB 2228 CGCCCTCTGGAGAGAACTTCGAACTGAGACTTCACACCTGAGCACTGAGAGA 2287
 QY 2281 GGAACCTCTGGAG 2340
 DB 2288 GGAACCTCTGGAG 2347
 QY 2341 CCAGAACTGCGGCTCACTATTTTGTATATAATGATGCTGAGCTGATCCAGCTGCTCGA 2400
 DB 2348 CCAGAACTGCGGCTCACTATTTTGTATATAATGATGCTGAGCTGATCCAGCTGCTCGA 2407
 QY 2401 CCTAC 2405
 DB 2408 CCTAC 2412

RESULT 5
 LOCUS BC009916 2430 bp mRNA linear PRI 12-NOV-2003
 DEFINITION Homo sapiens solute carrier family 27 (fatty acid transporter), member 3, mRNA (CDNA clone MGC:2739 IMAGE:2822875), complete cds.
 ACCESSION BC009916
 VERSION BC009916.2 GI:33869756
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2430)
 Strussberg, R.D., Collins, P.S., Wagner, L.H., Derge, J.G., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhac, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheer, T.E., Brownstein, M.J., Uedini, T.B., Toshiyuki, S., Carinini, P., Prange, C., Raha, S.S., Loguercio, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, J.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scheraga, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 2 (bases 1 to 2430)
 Strausberg, R.
 Direct Submission
 Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 On Aug 19, 2003, this sequence version replaced gi:14602825.
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINI)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadansystemsbiology.org
 Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

FEATURES
 source
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINI at: <http://image.llnl.gov>
 Series: IRAL Plate: 2 Row: P Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13236578.
 Location/Qualifiers
 1..2430
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="MGC:2739 IMAGE:2822875"
 /issue_type="Lung, small cell carcinoma"
 /clone_lib="NIH MGC 7"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..2430
 /gene="SLC27A3"
 /note="Synonyms: FATP3, MGC4365"
 /db_xref="LocustID:11000"
 /db_xref="MIM:604193"
 66..2258
 /product="solute carrier family 27 member 3"
 /protein_id="AA09916.1"
 /db_xref="GI:14602826"
 /db_xref="LocustID:11000"
 /translation="MGVQRTAPWKEKSOERALGFRKSGMFPAGMNPVPIE
 ASMAALLPLPLLPLLLKLHMLPOLRWLPALIAVVALCKRALRALAA
 ADPEEGGCSLAKRLAEAOBARHTLIGSRFSSEARESENRAARALV
 DMPPGGDSGEGSAGERRAAGADAAAGSAGSAGSAGSAGSAGSAGS
 LLLPAPRPMPLGLAKGLRTAPALRGLPLHCLRCGARRALTAPEFLE
 PLPLPRAVGLHMLNAGPOTHPAGISDLVAESAEDVPVPGYLSPOSITDTLYIF
 TSGTGLPRAARISHLKILOCQFYOLQVQEDVLYALPLNMSGSLAIVGMGI
 GATVVLKSFSAAGQFWECCOQRRVTFYIGELGRVNOPPSKARHKLVALVSG
 LRPDWKFSFVRSAGFLOVETGLTEGNVATINYGQAGVARSMLYKHLFPFELIR
 VYVTGERLRDQGCMAFSRPEGLVAAPVSOQSPFLGAVGPELAQGLKLVKVERP
 GVPNTGDLVCDQDQGLFRHRTDGRDTRFRKKGENVATTEVAEVEALDPLQENVVG
 QVPRHERRAGAAALVLRPPRALDMQLYTVSENLPRYAPRRRLQESLATETFK
 QQKRVANDEFPFSTLSPLVLDVAQVYVPLTTRARSLALAGLRV"
 71..1931
 /note="AMP-binding; Region: AMP-binding enzyme"
 /db_xref="CDD:pfam0501"

misc_feature
 ORIGIN
 Query Match 100.0%; Score 2405; DB 9; Length 2430;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACCTCTCCCGGCTTTCGCTCTCCGCCGCTGAGTGGAGGCGCTGGAGTAAT 60
 DB 8 GCACCTCTCCCGGCTTTCGCTCTCCGCCGCTGAGTGGAGGCGCTGGAGTAAT 67
 QY 61 GGGCGTGTGCCAGCGAGCGCGCTCCTCGAAGAGAAGTCTACGTAACAGCGCGC 120
 DB 68 GGGCGTGTGCCAGCGAGCGCGCTCCTCGAAGAGAAGTCTACGTAACAGCGCGC 127
 QY 121 CCTAGTTTTCGAGAGGAGATCAGGAGATGTTTCCGAGCGCGCTGGAACAGAGTGC 180
 DB 128 CTTAGTTTTCGAGAGGAGATCAGGAGATGTTTCCGAGCGCGCTGGAACAGAGTGC 187
 QY 181 GATAGAGAGAGCGGCTCCATGCTGCGCTCTCTGCTGCGCTCTGCTGTTTCTAC 240
 DB 188 GATAGAGAGAGCGGCTCCATGCTGCGCTCTCTGCTGCGCTCTGCTGTTTCTAC 247
 QY 241 GCTGCTGCTGAGAGCTACACCTCTGCGCGCAGTGGCGCTGCGCGCGAGCTGGC 300
 DB 248 GCTGCTGCTGAGAGCTACACCTCTGCGCGCAGTGGCGCTGCGCGCGAGCTGGC 307
 QY 301 CTTTTCGAGTGCAGCTCTGTCTGTGTAAGAGGAGCTCTTGTAGCTCGCGCGCGCGC 360
 DB 308 CTTTTCGAGTGCAGCTCTGTCTGTGTAAGAGGAGCTCTTGTAGCTCGCGCGCGCGC 367

QY	361	TTCCGCGGACCCGGAAAGTGTCCCAAGGGGGGCTGAGGCTGAGCTTGGCGGCTCCGGAAACT	420
Db	368	TGCGCGCGACCCGGAAAGTGTCCCAAGGGGGGCTGAGGCTGAGCTTGGCGGCTCCGGAAACT	427
QY	421	GGCCGACGACCGCGCGCGACACTTTCTCATTCAGGGCTCGCGGCGCTTTAGTACTC	480
Db	428	GGCCGACGACCGCGCGCGACACTTTCTCATTCAGGGCTCGCGGCGCTTTAGTACTC	487
QY	481	AGAGGCGGAGCGGAGGTAAACAGGGGTGCAACGCGCTTCTTACGTGCGCTTAGGCTGGGA	540
Db	488	AGAGGCGGAGCGGAGGTAAACAGGGGTGCAACGCGCTTCTTACGTGCGCTTAGGCTGGGA	547
QY	541	CTGGGGACCCGACGCGCGCGACAGCGCGGAGGGGAGCGCTGAGAAAGCGACGCGGACG	600
Db	548	CTGGGGACCCGACGCGCGCGACAGCGCGGAGGGGAGCGCTGAGAAAGCGGACGCGGACG	607
QY	601	GCCGGAGCCGGAGATGCAAGCGGCGGAAAGCGGCGGAGATTGGCCGAGGAGGACGAGTGC	660
Db	608	GCCGGAGCCGGAGATGCAAGCGGCGGAAAGCGGCGGAGATTGGCCGAGGAGGACGAGTGC	667
QY	661	CGCCAGAGGTGGAGAGAGCGCGCGCGCTGTACACTGAGAGCACTGTGGCGCGTGCCT	720
Db	668	CGCCAGAGGTGGAGAGAGCGCGCGCGCTGTACACTGAGAGCACTGTGGCGCGTGCCT	727
QY	721	CCCCGTGGGCCCAAGTTCTGTGAGCTCTGAGTTGGAGCTGGCCAAAGCCGCGCTGCGAC	780
Db	728	CCCCGTGGGCCCAAGTTCTGTGAGCTCTGAGTTGGAGCTGGCCAAAGCCGCGCTGCGAC	787
QY	781	TGCGCTTTGTGCGCCACCGCGCTTGCGCGCGGGGGCCCCCTGTGACCTGCACTGCGCAAGTGG	840
Db	788	TGCGCTTTGTGCGCCACCGCGCTTGCGCGCGGGGGCCCCCTGTGACCTGCACTGCGCAAGTGG	847
QY	841	CGCGCGCGCGCTGTGCTGCGCGCGCAAGTTCTGAGAGCCTCTGAGCGCGACCTGCGCGC	900
Db	848	CGCGCGCGCGCTGTGCTGCGCGCGCAAGTTCTGAGAGCCTCTGAGCGCGACCTGCGCGC	907
QY	901	CCTGAGAGCCATGAGGCTTCACTGTGAGCTTGCAAGGCCCAAGAACCACTCTGCTGAAAT	960
Db	908	CCTGAGAGCCATGAGGCTTCACTGTGAGCTTGCAAGGCCCAAGAACCACTCTGCTGAAAT	967
QY	961	TAGGAGTTTGTGCTGTGAAGTGTCCGCGTGAAGTGAATGGGCGACAGGACGATACCTCTC	1020
Db	968	TAGGAGTTTGTGCTGTGAAGTGTCCGCGTGAAGTGAATGGGCGACAGGACGATACCTCTC	1027
QY	1021	TTCCGCCCAAGCAATPACAGACAGGTGCTGTACATCTTCACTCTGCGACCAACGAGGCTT	1080
Db	1028	TTCCGCCCAAGCAATPACAGACAGGTGCTGTACATCTTCACTCTGCGACCAACGAGGCTT	1087
QY	1081	CCCCAAGGCTCTCGGATCACTGATCTTGAAGATCCTGGAATGCCAGGGCTTCTPACACT	1140
Db	1088	CCCCAAGGCTCTCGGATCACTGATCTTGAAGATCCTGGAATGCCAGGGCTTCTPACACT	1147
QY	1141	GTGTGATGTCCACACGAGAAAGTGTATCTTACCTCGCCCTTCCACTTACACATGTCCGG	1200
Db	1148	GTGTGATGTCCACACGAGAAAGTGTATCTTACCTCGCCCTTCCACTTACACATGTCCGG	1207
QY	1201	TTCCCTGCTGAGCATGTGGGCTGTGATGGGCAATGGGGCCACAGTGGTGTGAATCCAA	1260
Db	1208	TTCCCTGCTGAGCATGTGGGCTGTGATGGGCAATGGGGCCACAGTGGTGTGAATCCAA	1267
QY	1261	GTTCTCGGCTGATAGTTCTGGGAAAGTTGCGACAGACACAGGGTGAAGGTTTCCAGTA	1320
Db	1268	GTTCTCGGCTGATAGTTCTGGGAAAGTTGCGACAGACACAGGGTGAAGGTTTCCAGTA	1327
QY	1321	CATTGGGAGCTGTGCGGATACCTTTGTCAACAGTCCCGGAGCAAGGCAAGACGTGGCCA	1380
Db	1328	CATTGGGAGCTGTGCGGATACCTTTGTCAACAGTCCCGGAGCAAGGCAAGACGTGGCCA	1387
QY	1381	TAAAGTTCGAGCTGAGAGTGGGACAGCGGCTGCGCCACATAACTGAGAGCGTTTGTGCG	1440
Db	1388	TAAAGTTCGAGCTGAGAGTGGGACAGCGGCTGCGCCACATAACTGAGAGCGTTTGTGCG	1447
QY	1441	GCGCTTTCGGGCCCTTGCAAGTGTCTGAGACATATGGACTGACAGAGGCAACGTGGCCAC	1500

Db	1448	CGCGTTTCGGGCCCCCTGGAGGTTGCTGGAGACATATGGACTGACAGAGGGCAACGTCGGCAC	1507
QY	1501	CATCACTACTACACAGGACGACGGGGCCGCTGTGGGGCCGTGCTTCTTGAGCTTTACAAAGCATAT	1550
Db	1508	CATCACTACTACACAGGACGACGGGGCCGCTGTGGGGCCGTGCTTCTTGAGCTTTACAAAGCATAT	1567
QY	1561	CTTCCCCCTTCTCCCTGATTCGTGATGATGTCACACACAGAGAGCCAAATTCGGGACCCCCA	1620
Db	1568	CTTCCCCCTTCTCCCTGATTCGTGATGATGTCACACACAGAGAGCCAAATTCGGGACCCCCA	1627
QY	1621	GGGGCACTGATGAGCCACATCTTCCAGGTGAGCCAGGGCTCTGTGTGGCCCCCGTAAACCA	1680
Db	1628	GGGGCACTGATGAGCCACATCTTCCAGGTGAGCCAGGGCTCTGTGTGGCCCCCGTAAAGCCA	1687
QY	1681	GCAGTCCCCAATTCCTGGGCTATGCTGTGCGGGGCCAGAGCTGAGCCCAAGGGAAAGTTGCTAA	1740
Db	1688	GCAGTCCCCAATTCCTGGGCTATGCTGTGCGGGGCCAGAGCTGAGCCCAAGGGAAAGTTGCTAA	1747
QY	1741	GGAGTCTTTCGGGCTTGGGAGTGTTCCTTCAACACTGGGGAACCTGTGTGCTGCATGA	1800
Db	1748	GGAGTCTTTCGGGCTTGGGAGTGTTCCTTCAACACTGGGGAACCTGTGTGCTGCATGA	1807
QY	1801	CCAAAGTTTTCCTCGGCTTCATGATGCTACTGAGAGAACCTTTCAGGTGGAAAGGGGGGAGA	1860
Db	1808	CCAAAGTTTTCCTCGGCTTCATGATGCTACTGAGAGAACCTTTCAGGTGGAAAGGGGGGAGA	1867
QY	1861	TGTGGCCACACACGAGGTGGCAGAGGTCTTTCGAGGCCCTGATATTTCTTACAGAGGTGA	1920
Db	1868	TGTGGCCACACACGAGGTGGCAGAGGTCTTTCGAGGCCCTGATATTTCTTACAGAGGTGA	1927
QY	1921	CGTCTATGAGATCATGTCGTGCCAGGGCATGAAAGCAGGGCTGGAAATGGCAGCCTTAGTTCT	1980
Db	1928	CGTCTATGAGATCATGTCGTGCCAGGGCATGAAAGCAGGGCTGGAAATGGCAGCCTTAGTTCT	1987
QY	1981	GGGTCCCCCCCCAGGCTTTGACCTTATGACAGCTTACACCACAGTGTCTGAGAACTTGGC	2040
Db	1988	GGGTCCCCCCCCAGGCTTTGACCTTATGACAGCTTACACCACAGTGTCTGAGAACTTGGC	2047
QY	2041	ACCTTATGCCCGGCCCCCGATTCCTCAGAGCTTCAGAGATCTTTGGCCAACACAGAGACCTT	2100
Db	2048	ACCTTATGCCCGGCCCCCGATTCCTCAGAGCTTCAGAGATCTTTGGCCAACACAGAGACCTT	2107
QY	2101	CAAAACAGCAAAAGTTGGATGGCAATGAGAGGCTTGAACCCACAGCAACCTGTGTGACCC	2166
Db	2108	CAAAACAGCAAAAGTTGGATGGCAATGAGAGGCTTGAACCCACAGCAACCTGTGTGACCC	2167
QY	2161	ACTGTACGTTTGGACACAGGCTGTAGGTGCTTACCTGCCCTTCACAACTGCGCGGTACAG	2220
Db	2168	ACTGTACGTTTGGACACAGGCTGTAGGTGCTTACCTGCCCTTCACAACTGCGCGGTACAG	2227
QY	2221	CGCCCTCTTGGACAGAAACCTTGCATCTTGAAACCTTCCACACTGAGAGCACTGAGAGA	2280
Db	2228	CGCCCTCTTGGACAGAAACCTTGCATCTTGAAACCTTCCACACTGAGAGCACTGAGAGA	2287
QY	2281	GGAGTCTGTGTGGGAGGGGGGCGGCTGTGACAGTGTACTGGGCTGTGAGGAGATCTTTCTATA	2340
Db	2288	GGAGTCTGTGTGGGAGGGGGGCGGCTGTGACAGTGTACTGGGCTGTGAGGAGATCTTTCTATA	2347
QY	2341	CCAGAACTGCGGTCACATATTTTGTAAATAATGTGAGCTGAGACTGATCCAGCTGTCTCTGA	2400
Db	2348	CCAGAACTGCGGTCACATATTTTGTAAATAATGTGAGCTGAGACTGATCCAGCTGTCTCTGA	2407
QY	2401	CTTAC 2405	
Db	2408	CTTAC 2412	

RESULT 6		
AXI01424		
LOCUS		
DEFINITION	Sequence 116 from Patent WO0121795.	DNA
ACCESSION	AXI01424	linear
PAT 10-APR-2001		

linear PAT 10-APR-2001

Db	1561	TCTCCTTATTTGCTATATGTCAACAAGAGAGCCAAATTCGGAGCCCCCAGGGGCACT	1620
OY	1629	GTAATGGCAACATCTCCAGGTGAAGCCAGGGCTGTGTGTGGCCCCGGTAAGCAGCAGTCCC	1688
Db	1621	GTATGGCCCAATCTCCAGGTGAAGCCAGGGCTGTGTGTGGCCCCGGTAAGCCAGCAGTCCC	1680
OY	1689	CATTCTGTGGGCTATAGCTGGGGGGCCAGAGGCTGGGCCAAGGGGAAGTGTGTAAGAATGCT	1748
Db	1681	CATTCTGTGGGCTATAGCTGGGGGGCCAGAGGCTGGGCCAAGGGGAAGTGTGTAAGAATGCT	1740
OY	1749	TCCGGCCTGGGGAATGTTTTCTTCAACACTGGGGAACTGTGCTGTCCGATGACCAAGTT	1808
Db	1741	TCCGGCCTGGGGAATGTTTTCTTCAACACTGGGGAACTGTGCTGTCCGATGACCAAGTT	1800
OY	1809	TTCTTCGGCTTCCATATATCTGTAACAGACCTTCAAGTGGAAAGGGGGAGAAATGTGGCA	1868
Db	1801	TTCTTCGGCTTCCATATATCTGTAACAGACCTTCAAGTGGAAAGGGGGAGAAATGTGGCA	1860
OY	1869	CAACCGAAGTGGCAGAGGCTCTTCGAGGCGCTAGATTTTCTTCAGAGGCTGACGTCTATG	1928
Db	1861	CAACCGAAGTGGCAGAGGCTCTTCGAGGCGCTAGATTTTCTTCAGAGGCTGACGTCTATG	1920
OY	1929	GAGTCACTGTGCCAGGGGATGAAGGCGAGGCTGGAATGGCAGGCTTAATTCTGGGTCCCC	1988
Db	1921	GAGTCACTGTGCCAGGGGATGAAGGCGAGGCTGGAATGGCAGGCTTAATTCTGGGTCCCC	1980
OY	1989	CCCAACGCTTTGAGACCTTATGACAGCTCTACACCAAGTGTGAGAGAACTTGCCACCTTATG	2048
Db	1981	CCCAACGCTTTGAGACCTTATGACAGCTCTACACCAAGTGTGAGAGAACTTGCCACCTTATG	2040
OY	2049	CCCGGCCCCGATTTCTCAAGGCTCCAGAGGCTTTTGGCCACACAGAGACCTTCAACAGC	2108
Db	2041	CCCGGCCCCGATTTCTCAAGGCTCCAGAGGCTTTTGGCCACACAGAGACCTTCAACAGC	2100
OY	2109	AGAAAGTTCCGATGGGCAATGAGAGGGCTTCGAGCCCGAGACCTGTGCTGAACCACTGATG	2168
Db	2101	AGAAAGTTCCGATGGGCAATGAGAGGGCTTCGAGCCCGAGACCTGTGCTGAACCACTGATG	2160
OY	2169	TTCTTGACACAGGCTGTAGTGTGCTACCTGTCCCTTCACAACCTGCCGGATACAGCGCCCTC	2228
Db	2161	TTCTTGACACAGGCTGTAGTGTGCTACCTGTCCCTTCACAACCTGCCGGATACAGCGCCCTC	2220
OY	2229	TGGCAGGAAGACCTTCGAACTTGAGAACTTTCACACCTGAAGGACCTGTGAGAGAGAACTCT	2288
Db	2221	TGGCAGGAAGACCTTCGAACTTGAGAACTTTCACACCTGAAGGACCTGTGAGAGAGAACTCT	2280
OY	2289	GTGGGGTGGGGCCGTTGACAGTGTACTGGGCTGTCAAGGAACTTTTCTATACAGAACT	2348
Db	2281	GTGGGGTGGGGCCGTTGACAGTGTACTGGGCTGTCAAGGAACTTTTCTATACAGAACT	2340
OY	2349	GCGGTCATATATTTTGTATAATATGTGGCTGAGAGCTGATCCAGGCTCTCTCTGA	2400
Db	2341	GCGGTCATATATTTTGTATAATATGTGGCTGAGAGCTGATCCAGGCTCTCTCTGA	2392

RESULT 7	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AY358409	AY358409	Human sapiens clone	DNA50913	2414 bp	mRNA	linear	PRI 03-OCT-2003		
			55C827A3	(UNQ367)	mRNA	complete cds.			
			AY358409.1	GI:37181942					
			FLI-CDNA.						
			Human sapiens	(human)					
			Human sapiens						
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
			Mammalia; Eucnemia; Primates; Catarrhini; Homnidae; Homo.						
			1 (bases 1 to 2414)						
			Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,						

Chen, J., Chow, B., Chui, C., Crowley, C., Currell, B., Deuel, B., Dowd, P., Eaton, D., Foster, J., Grimaldi, C., Gu, Q., Hass, P. E., Heldens, S., Huang, A., Kim, H. S., Klimowski, L. M., Jin, Y., Johnson, S., Lee, J., Lewis, L., Liao, D., Mark, M., Robbie, E., Sanchez, C.,

TITLE	Schoenfeld, J., Seshagiri, S., Simmons, L., Singh, J., Smith, V., Stinson, J., Vagstad, A., Vandlen, R., Watanabe, C., Weidner, D., Woods, K., Xie, M.-H., Yasuda, D., Yi, S., Yu, G., Yuan, Y., Zhang, M., Zhang, Z., Goddard, A., Wood, M. I. and Godowski, P.					
JOURNAL	The Secreted Protein Discovery Initiative (SPDI), a Large-Scale Effort to Identify Novel Human Secreted and Transmembrane Proteins: A Bioinformatics Assessment					
REFERENCE	Genome Res. 13 (10), 2265-2270 (2003)					
AUTHORS	12975309					
TITLE	2 (bases 1 to 2414)					
JOURNAL	Clark, H. F.					
FEATURES	Direct Submission Submitted (01-AUG-2003) Department of Biotechnology, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA					
Source	Location/Qualifiers					
gene	1..2414					
CDS	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="DNA50913" 1..2414 /locus_tag="UNQ367" 54..2246 /locus_tag="UNQ367" /note="PRO703" /codon_start=1 /product="SLC27A3" /protein_id="PAC088775.1" /db_xref="GI:37181943" /translation="MGVCQRTAPWKSGLERALLFRKSGSMFASGMNQYPIEEAGSNALLPLPLPLLLPLLLKLHWLQLMPLPDLLFAVALDCKRALPARLAAADPEPGEGSGSLAMRLAEALAOQRAMTFILHGSRFSYSABERSNSNAAFLEALWDMDGSDGGSEGSAGEBERAAPAGDAAGSGAEFLAGDGAARGGAAPASPGATVVAIPLIPAGEFLFWLMFGAKAGRIFARFPAFTLRSGFLHLCLSCGRALVLAPELFSELPALPAPAMGLHMLANAGRPAGISPLALREVSADVDPVPYRHSPOSTIDTCLYLITSGCTGTAKRAAISHLKILOCGGFYOICGHOPVLYIALPYMSGSLIGVCMGIGATVLSKRSAGCFMEDCOQHRYTVFCYIGELCRYLVNPPSKALEHKYRLAVSGGLRPDWTERFVRFRFEPQLVLETYGTESENVAITINTGORGAVGRASMLYKHLPSPSLIRYDVYTFEPIRDPOCHMATSPGEBSGLIVAPVSQSPGLGYAAGVELAQKLIKQVFRRGDVEFNTGDLVLCDDQGLFHFHDRTCDTPFMKGENTVAETBAEVFEALDPLEOVNVVGVVYVGEHRAGRAMALVLRPHALDLMVLTHVSNLDPVAPARPRLROESLATETEFKKQOKRMANEGFDSTLSDPLYLDQAAGAVLPLTTARYSALLAENLI"					
ORIGIN						
Query Match	99.3%; Score 2387; DB 9; Length 2414;					
Best Local Similarity	100.0%; Pired. No. 0;					
Matches 2387; Conservative	0; Mismatches 0; Indels 0; Gaps 0;					
Oy	18	CTGCTCCCGCGCGTGGAAGTGGTGGAGGGGCTGGTGGGAATGGAGGCTGGCCAGCGCA	77			
Dd	13	CTGCTCTCCGCCCTGTGGAGTGGTGGAGGGGCTGGTGGGAATGGAGGCTGGCCAGCGCA	72			
Oy	78	CGCGCGCTCCCTGGAGAAGTAGTCAGCTAGAACAGCGGCCCTAAGTTTTCCGAAGG	137			
Dd	73	CGCGCGCTCCCTGGAGAAGTAGTCAGCTAGAACAGCGGCCCTAAGTTTTCCGAAGG	132			
Oy	138	GAGANTCAGGAGATTGTGGCAGCGCGCTGGAAACCAGAGGTGCCCATAGAGAACCGGGCT	197			
Dd	133	GAGANTCAGGAGATTGTGGCAGCGCGCTGGAAACCAGAGGTGCCCATAGAGAACCGGGCT	192			
Oy	198	CCATGGCTGCCTCTCTCTGCTGCCTCTGCTGCTGTGTACTACCCCTGCTGCTGTGAAGC	257			
Dd	193	CCAATGGCTGCCTCTCTCTGCTGCCTCTGCTGCTGTGTACTACCCCTGCTGCTGTGAAGC	252			
Oy	258	TACACCTCTGGCCGCAAGTTGCGCTGGCTTCCGGCGGAATTGACCTTTTGCGTGGAGCTC	317			
Dd	253	TACACCTCTGGCCGCAAGTTGCGCTGGCTTCCGGCGGAATTGACCTTTTGCGTGGAGCTC	312			
Oy	318	TGTGCTGCAAAAAGGGCTCTTGAGACTCGCGCCCTGGCCGCGAGTCCGACCCGGAAG	377			
Dd	313	TGTGCTGCAAAAAGGGCTCTTGAGACTCGCGCCCTGGCCGCGAGTCCGACCCGGAAG	372			
Oy	378	GTCGCCAGGGGGGCTGCAGCGCTGGCTTGCGGCTTCGCGGAATGGCCACGACGCGCGC	437			

QY 1938 TCCAGAGGATGATGAGGAGGAGGCTGGAAATGAGCCCTGATGCTGCTCCGCCACGCTT 1997
 DB 1994 TCCAGAGGATGATGAGGAGGAGGCTGGAAATGAGCCCTGATGCTGCTCCGCCACGCTT 2053
 QY 1998 TCGACCTTATGACAGCTCTACACCCAGCTGTCTGAGAACTTGGCACTTATGCTCCGCCCTC 2057
 DB 2054 TCGACCTTATGACAGCTCTACACCCAGCTGTCTGAGAACTTGGCACTTATGCTCCGCCCTC 2113
 QY 2058 GATTCCTGAGGCTCCAGGAGGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2117
 DB 2114 GATTCCTGAGGCTCCAGGAGGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2173
 QY 2118 GGATGCGAAATGAGGAGGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2177
 DB 2174 GGATGCGAAATGAGGAGGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2233
 QY 2178 AGGCTGTGAGGCTCTACAGGAGGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2237
 DB 2234 AGGCTGTGAGGCTCTACAGGAGGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2293
 QY 2238 ACCTTCGAACTGAGAACTTCCACACCTGAGGAGCTTGAAGAGAGAACTTGTGGGCTGG 2297
 DB 2294 ACCTTCGAACTGAGAACTTCCACACCTGAGGAGCTTGAAGAGAGAACTTGTGGGCTGG 2353
 QY 2298 GGGCGCTTGGAGGCTTGTGAGGAGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2357
 DB 2354 GGGCGCTTGGAGGCTTGTGAGGAGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2413
 QY 2358 ATTTGTATATATGAGGCTGAGGAGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2404
 DB 2414 ATTTGTATATATGAGGCTGAGGAGCTTGTGGCCACAGAGAGCTTCAAGAGAGAAAGTTC 2460

RESULT 9
 AK125102 3343 bp mRNA linear PRI 09-SEP-2003
 LOCUS Homo sapiens CDNA FLJ43112 fls, clone CTONG2028124, weakly similar
 DEFINITION to Homo sapiens VLCS-H1 protein (VLCS-H1).
 ACCESSION AK125102.1 GI:34531083
 VERSION oligo capping; fls (full insert sequence).
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S.,
 Tanei, H., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Watanabe, M.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, U., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
 Kimura, K., Matsuo, K., Nakamura, Y., Sakine, M., Kikuchi, H., Kanda, K.,
 Wagahtsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A.,
 Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K.
 and Isegaki, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 3343)
 Isegaki, T. and Yamamoto, J.
 Direct Submission
 Submitted (15-JUN-2003) Takao Isegaki, Flj Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kiserazu, Chiba 292-0818, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB) (supported by Japan
 Construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; Clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1..3343
 /organism="Homo sapiens"

FEATURES
 source

ORIGIN
 Query Match 99.3%; Score 2387; DB 9; Length 3343;
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 1; Gaps 1;
 Matches 2398; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 1 GCACCTCTCCGCGGATTTGCTCTCCGCGGCTGAGAGAGTGTGGGCGCTGGGAGT 60
 DB 946 GCACCTCTCCGCGGATTTGCTCTCCGCGGCTGAGAGAGTGTGGGCGCTGGGAGT 1005
 61 GGGCGCTGTGCGAGCGGAGCGGCTTCCCTGAGAGAGAGTCTGACCTGAGAGCGGC 120
 DB 1006 GGGCGCTGTGCGAGCGGAGCGGCTTCCCTGAGAGAGAGTCTGACCTGAGAGCGGC 1065
 121 CTTAGGTTTCCGAGAGAGAGTGTGTTGGAGCGGCTGAGAGCGGCTGAGAGCGGC 180
 DB 1066 CTTAGGTTTCCGAGAGAGAGTGTGTTGGAGCGGCTGAGAGCGGCTGAGAGCGGC 1125
 181 GATAGAGAGAGCGGCTGAGAGCGGCTTCCCTGAGAGAGAGTCTGACCTGAGAGCGGC 240
 DB 1126 GATAGAGAGAGCGGCTGAGAGCGGCTTCCCTGAGAGAGAGTCTGACCTGAGAGCGGC 1185
 241 GCTGCTGCTGCTGAGAGTCACTGCTGAGCGGAGTGGCTGAGAGCGGCTGAGAGCGGC 300
 DB 1186 GCTGCTGCTGCTGAGAGTCACTGCTGAGCGGAGTGGCTGAGAGCGGCTGAGAGCGGC 1245
 301 CTTTGGGCTGCGAGCTCTGCTGAGAGAGGCTTCTGAGAGCGGCTGAGAGCGGC 360
 DB 1246 CTTTGGGCTGCGAGCTCTGCTGAGAGAGGCTTCTGAGAGCGGCTGAGAGCGGC 1305
 361 TGGCGCGGAGCGGAGAGGCTTCCGAGAGGAGGCTGAGAGCGGCTGAGAGCGGC 420
 DB 1306 TGGCGCGGAGCGGAGAGGCTTCCGAGAGGAGGCTGAGAGCGGCTGAGAGCGGC 1365
 421 GGGCGAGAGCGGCGGAGAGGCTTCTGAGAGAGGCTTCTGAGAGCGGCTGAGAGCGGC 480
 DB 1366 GGGCGAGAGCGGCGGAGAGGCTTCTGAGAGAGGCTTCTGAGAGCGGCTGAGAGCGGC 1425
 481 AGAGCGGAGCGGAGAGGCTTCTGAGAGAGGCTTCTGAGAGCGGCTGAGAGCGGC 540
 DB 1426 AGAGCGGAGCGGAGAGGCTTCTGAGAGAGGCTTCTGAGAGCGGCTGAGAGCGGC 1485
 541 CTTGGGAGCGGAGCGGAGAGGCTTCTGAGAGAGGCTTCTGAGAGCGGCTGAGAGCGGC 600
 DB 1486 CTTGGGAGCGGAGCGGAGAGGCTTCTGAGAGAGGCTTCTGAGAGCGGCTGAGAGCGGC 1545
 601 GCCGAGCGGAGAGATGAGCGGCGGAGAGCGGCGGAGAGTTCGCGAGGAGAGCGGTC 660
 DB 1546 GCCGAGCGGAGAGATGAGCGGCGGAGAGCGGCGGAGAGTTCGCGAGGAGAGCGGTC 1605

CDS

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CTONG2028124"
 /tissue_type="concrete, tumor tissue"
 /clone_id="CTONG2"
 /note="cloning vector: pME185FLJ3"
 761..2683
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC86050.1"
 /db_xref="GI:34531084"
 /translation="MPKPRPVNLEDRYNGIQRREHPPGSRVPAASSPPGSAAP
 AERRLGGMBESGAAPHPHSPRVNLEDRYNGIQRREHPPGSRVPAASSPPGSAAP
 LGRKGGSGVFASGMNQVPIEASMAALLLPLALPLLELLKMLPQRLPAD
 LAFAYALCCXKALRALALAAADPEEGGCSLAWRLAEALQQAATFLIHSRR
 FSYSEARESNRAARAFILALGMDWGPDDGSGESGASGERAPAGADAAAGSAEF
 AGGDSAGRGGAAPLSPGATVALLPAPPEVPLIMFGLAKGLTAAPVPLARGLP
 LHCRGCGARALVIAPEPLESPEPLPALRAMGLHMAAGPETHAGISDLAEVSAE
 VDGPRVGYLSRPSQITTDCLVYFTSGCTGLPRAARISHLKLCCGFPYOLGQVHEDV
 IYIALPLYHSSSLGIIVCMGIGATVILKSFSAQFPREDQQRVTVFQYIGELCR
 YLVNDPPPSFAERGHKRVRLVSGLEPDTWERVRRFGPQLVETIGLBNVATINTY
 GORGAVGRASMPYKHIFPESLIRYDVTGTGEPIRDPQHCMAVSPPEPGLVAVASSP
 HSWAMLAGQSWPRGSC"

QY	661	CGCCAGAGCTGAGAGAGCGCGCGCCCTCTGTGACCTGAGACATGTGAGGCGTCTCT	720
Dp	1606	CGCCAGAGCTGAGAGAGCGCGCGCCCTCTGTGACCTGAGACATGTGAGGCGTCTCT	1665
QY	721	CCCCGCTGAGCCAGATTCTGTGGCTGTGGTTGGGCTGGCCAGAGCGGAGCTGCGAC	780
Dp	1666	CCCCGCTGAGCCAGATTCTGTGGCTGTGGTTGGGCTGGCCAGAGCGGAGCTGCGAC	1725
QY	781	TGCGTTGTGGCCACCGGCCGTGGCGCGGGGGCCCTGCTGACGTGCTCCGCAAGTGGG	840
Dp	1726	TGCGTTGTGGCCACCGGCCGTGGCGCGGGGGCCCTGCTGACGTGCTCCGCAAGTGGG	1785
QY	841	CGCGCGCGCGGTGTGTGTGGCGCCAGATTCTGGAGTCCCTGAGCGGACCTGCGC	900
Dp	1786	CGCGCGCGCGGTGTGTGTGGCGCCAGATTCTGGAGTCCCTGAGCGGAGCTGCGC	1845
QY	901	CCTGAGACCATGGGGCTCCACCTGTGGGCTGACAGGCCACAGAACCCACCTGCTGAA	960
Dp	1846	CCTGAGACCATGGGGCTCCACCTGTGGGCTGACAGGCCACAGAACCCACCTGCTGAA	1905
QY	961	TAGGCATTGTGCGCTGAAGTGTCCGCTGGAAGTGAATGGGCGACGAGGATACCTTC	1020
Dp	1966	TAGGCATTGTGCGCTGAAGTGTCCGCTGGAAGTGAATGGGCGACGAGGATACCTTC	1965
QY	1021	TTCCCCCAGAGCATACAGACAGGTGCTGTACATCTTCACTCTGACACAGGAGCT	1080
Dp	1966	TTCCCCCAGAGCATACAGACAGGTGCTGTACATCTTCACTCTGACACAGGAGCT	2025
QY	1081	CCCCAAGGCTGCTGGATCAGTCACTGTAABATCCGTGAAGTCCAGAGGCTCTATACGT	1140
Dp	2026	CCCCAAGGCTGCTGGATCAGTCACTGTAABATCCGTGAAGTCCAGAGGCTCTATACGT	2085
QY	1141	GTTGTGTGTCCACAGAGAAATGTGTATCTACCTCGCTCCCATCTTACACATGTCCG	1200
Dp	2086	GTTGTGTGTCCACAGAGAAATGTGTATCTACCTCGCTCCCATCTTACACATGTCCG	2145
QY	1201	TTCCCTGTGGGACATCTGTGGCTGTGATGGGCATTGGGGCCACAGTGTGTGAAATCCA	1260
Dp	2146	TTCCCTGTGGGACATCTGTGGCTGTGATGGGCATTGGGGCCACAGTGTGTGAAATCCA	2205
QY	1261	GTTCTCGGCTGATCACTCTGGGAAAGTTGCGACAGACAGAGGGTGAAGGTGTCCAGTA	1320
Dp	2266	GTTCTCGGCTGATCACTCTGGGAAAGTTGCGACAGACAGAGGGTGAAGGTGTCCAGTA	2265
QY	1321	CATTGGGAGCTGTGCGGATACCTTGTCAACCAAGCCCGCGAGCAAGGACGATGGCCA	1380
Dp	2266	CATTGGGAGCTGTGCGGATACCTTGTCAACCAAGCCCGCGAGCAAGGACGATGGCCA	2325
QY	1381	TAAAGTCCGCGTGGAGTGGGAGGAGGGCTGGGCCATATCCTGGGAGCGTTTGTGGG	1440
Dp	2326	TAAAGTCCGCGTGGAGTGGGAGGAGGGCTGGGCCATATCCTGGGAGCGTTTGTGGG	2385
QY	1441	GCGCTTGGGGCCCTGACAGTGTCTGAGACATATGACTGAACAGGGCAACGTGGCCAC	1500
Dp	2386	GCGCTTGGGGCCCTGACAGTGTCTGAGACATATGACTGAACAGGGCAACGTGGCCAC	2445
QY	1501	CATCAACATCACAGGACAGCGGGGCGCTGTGGGGGTCTTCTGTGGCTTTACAGCAAT	1560
Dp	2446	CATCAACATCACAGGACAGCGGGGCGCTGTGGGGGTCTTCTGTGGCTTTACAGCAAT	2505
QY	1561	CTTCCCTCTTCTCTGATTGTGTATGATGATCACACACAGAGAGCCAAATTCGGAGC	1620
Dp	2506	CTTCCCTCTTCTCTGATTGTGTATGATGATCACACACAGAGAGCCAAATTCGGAGC	2565
QY	1621	GGGGACCTGTATGGCCATCTCCAGGTGAGCCAGGGCTGCTGTGTGGTCCCGGTAAAGCCA	1680
Dp	2566	GGGGACCTGTATGGCCATCTCCAGGTGAGCCAGGGCTGCTGTGTGGTCCCGGTAAAGCCA	2624
QY	1681	GCAGTCCCATCTTCTGGGCTATGTCTGGCGGAGCAGAGCTGGGCCACAGGGGAATGTCTAA	1740
Dp	2625	GCAGTCCCATCTTCTGGGCTATGTCTGGCGGAGCAGAGCTGGGCCACAGGGGAATGTCTAA	2684
QY	1741	GGATGTCTTCCGCGCTGGGATGTTTCTTCAACCTGGGACCTGTGTCTGCGATGA	1800

Db	2685	GGATGCTTCGCGCGCTGGGGATGTTTTCTTCACACACTGGGGACCTGCTGCTCGAATGA	2744
QY	1801	CCAAAGTTTTCTCCGCTTCCATGATGCTACTGAGACACCTTCAGGTGGAAAGGGGAGAA	1866
Db	2745	CCAAAGTTTTCTCCGCTTCCATGATGCTACTGAGACACCTTCAGGTGGAAAGGGGAGAA	2804
QY	1861	TGATGGCCACAACCGAGGTGGCAGAGGTCTTCAGAGCCCTAGATTTTCTTCAGAGAGTAA	1920
Db	2805	TGATGGCCACAACCGAGGTGGCAGAGGTCTTCAGAGCCCTAGATTTTCTTCAGAGAGTAA	2866
QY	1921	CGTCTATGAGTCACTGTGCCAGGGCATGAAGCGAGGGCTGGAAATGCGAGCCCTTAGTCT	1980
Db	2865	CGTCTATGAGTCACTGTGCCAGGGCATGAAGCGAGGGCTGGAAATGCGAGCCCTTAGTCT	2924
QY	1981	GGCGTCCCCCACAAGCTTTGACCTTATGACAGCTTACACCAAGTGTCTGGAACCTTGGC	2040
Db	2925	GGCGTCCCCCACAAGCTTTGACCTTATGACAGCTTACACCAAGTGTCTGGAACCTTGGC	2984
QY	2041	ACCTTATGCCCCGCCCCGATTTCTCAGGCTTCCAGAGTCTTTGGCCACCAAGAGACCTT	2100
Db	2985	ACCTTATGCCCCGCCCCGATTTCTCAGAGCTTCCAGAGTCTTTGGCCACCAAGAGACCTT	3044
QY	2101	CAAAACACAGAAAGTTGGATGGCAATGAGGGCTTGAACCCAGACACCCGTGTGACCC	2160
Db	3045	CAAAACACAGAAAGTTGGATGGCAATGAGGGCTTGAACCCAGACACCCGTGTGACCC	3104
QY	2161	ACTGTAGCTTCTGAGCACAGGCTGTAGTGGCTTACCTGCCCTTCACAACTGCCCGGTACG	2220
Db	3105	ACTGTAGCTTCTGAGCACAGGCTGTAGTGGCTTACCTGCCCTTCACAACTGCCCGGTACG	3164
QY	2221	CGCCCTCTGGCAGGAACGTTGGAATCTGAGACCTTCACACACTGAGGGACCTGAGAGA	2280
Db	3165	CGCCCTCTGGCAGGAACGTTGGAATCTGAGACCTTCACACACTGAGGGACCTGAGAGA	3224
QY	2281	GGAATCTGTGGGGTGGGGGCCGTTGTAGAGTGTACTGGGCTGTACAGGAGATCTTTTCTATA	2340
Db	3225	GGAATCTGTGGGGTGGGGGCCGTTGTAGAGTGTACTGGGCTGTACAGGAGATCTTTTCTATA	3284
QY	2341	CCAGAACTGGGGTCACTAATTTTGTAAATAATGTGGCTGGAGCGTATCCAGCTGTCTCTCG	2399
Db	3285	CCAGAACTGGGGTCACTAATTTTGTAAATAATGTGGCTGGAGCGTATCCAGCTGTCTCTCG	3343

RESULT	10
LOCUS	AX882240
DEFINITION	Sequence 17145 from Patent EP1074617.
ACCESSION	AX882240
VERSION	AX882240.1
KEYWORDS	GI:40037055
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	Oca,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T. Primers for synthesising full-length cDNA and their use Patent: EP 1074617-A 17145 07-FEB-2001;
JOURNAL	Research Association for Biotechnology (JP) Location/Qualifiers 1..2314
FEATURES	
source	

```

CDS
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        60..2162
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="CAE91498.1"
        /db_xref="GI:40037056"
        /translation="MFASGNQIVPTEAGSVALLLPILLPILLKLIHPQIR
        WLPADIAFAVRLCCKALRALAAAPPEBEGGCSIAMRLAEIAGQRAATFTII

```

ORIGIN

Query Match 96.1%; Score 2310.8; DB 6; Length 2314;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2312; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

HGRSRFSYSEARESENRARAFALFALGMDWDPDGDGSGSGSAGEERRAPGADPAAG
 SAAEPAGDGGARGGGMAAPLSRGAITVALLPAGPERPMLRGLAKGRTAPRPTLL
 RRPFLHCRSCGARALVAPBLESLEBDLPALBAPLHMAPBHPHGISDLA
 EVASVDEPVPGLSSPOSITDTCLYITPSGTGLPKARISHKLIQCGFYOLGV
 HOEDVLYALPLVHMSGSLIGVCMGIGATVLSKFSQAQFEDCOQHVTFQYI
 GELCRVLYVQPSKAREGHKVLAVGSLRPTWERFVRREGPLQVETTYLTGNVA
 TINYTGORGAVRASMLYKHIPFSLIHYDTGTGPIDPOGCMATSPGRLVAV
 VSOOSPLFVAGAPBELLAOGLIKIDVFRGADVFENFGDLVDDGSLFPHRTGDTFR
 AKGVNATTEVAEVEFALDPLDEVNVVYVTPVGHGHRGAMALYLRPHALDMQIXT
 HVENLPPYARPRFLRLOSLATITETFRQOKVRMANBEPDSTLSDPULVDQAVAGI
 LPLTARYSALLAGNLRIT
 90 GGAAGAGAAAGTCTAGCTAGAACAGACGGCCCTTAGCTTTTCGAGAGAGATCAGGGA 149
 1 GGAAGAGAAAGTCTAGCTAGAACAGACGGCCCTTAGCTTTTCGAGAGAGATCAGGGA 60
 150 TGTTCGCGAGCGCTGGAAACAGACGGTCCGATAGAGAAAGCGGGCTCCATGGCTGCC 209
 61 TTTTTCGACGCGGCTGGAAACAGACGGTCCGATAGAGAAAGCGGGCTCCATGGCTGCC 120
 210 TCTTCTGCTGCCCCCTGCTGCTGTTGCTAACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 269
 121 TCTTCTGCTGCCCCCTGCTGCTGTTGCTAACCGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 270 CGCAGTTGGCGTGGCTTCGCGGAGACTTGGCCCTTGGCGGTGGAGCTGTGTGCTGAGAA 329
 181 CGAGTTGGCGTGGCTTCGCGGAGACTTGGCCCTTGGCGGTGGAGCTGTGTGCTGAGAA 240
 330 GGGCTCTTTGAGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCG 389
 241 GGGCTCTTTGAGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCG 300
 390 GGTTCAGCTTGGCTTCGCGGAGACTTGGCCCTTGGCGGTGGAGCTGTGTGCTGAGAA 449
 301 GGTTCAGCTTGGCTTCGCGGAGACTTGGCCCTTGGCGGTGGAGCTGTGTGCTGAGAA 360
 450 TCAATTCAGCTTCGCGGCTTCGCGGAGACTTGGCCCTTGGCGGTGGAGCTGTGTGCTGAG 509
 361 TCAATTCAGCTTCGCGGCTTCGCGGAGACTTGGCCCTTGGCGGTGGAGCTGTGTGCTGAG 420
 510 CACGCGCTTCTTCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 569
 421 CACGCGCTTCTTCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 570 AGGAGAGCGCTGAGAGAGCGAGCGGAGCGAGCGGAGCGGAGATCAGCGGCGGGA 629
 481 AGGAGAGCGCTGAGAGAGCGAGCGGAGCGAGCGGAGCGGAGATCAGCGGCGGGA 540
 630 GCGGCGCGGAGTTTTCGCGAGGAGAGCGGCTCGCGGAGAGCTGAGAGAGCGGCGGCTC 689
 541 GCGGCGCGGAGTTTTCGCGAGGAGAGCGGCTCGCGGAGAGCTGAGAGAGCGGCGGCTC 600
 690 TGTTCAGCTTCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
 601 TGTTCAGCTTCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 750 GGTTCGCGCTGCGCAAGCGCGGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809
 661 GGTTCGCGCTGCGCAAGCGCGGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 810 GCGGCGCTGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869
 721 GCGGCGCTGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 870 TTTTCGAGAGCTTCGAGAGCGGAGCTGCGGCTTCGAGAGAGATGAGGCTTCACCTGTGGG 929
 781 TTTTCGAGAGCTTCGAGAGCGGAGCTGCGGCTTCGAGAGAGATGAGGCTTCACCTGTGGG 840

QY CTGACGGCCCGAGGAGCCACCCCTGCTGATTAAGCATTTGTGCTGCTGAGTGTCCGCTG 989
 DB CTGACGGCCCGAGGAGCCACCCCTGCTGATTAAGCATTTGTGCTGCTGAGTGTCCGCTG 900
 QY AAGTGAATGGGCGCAGTGGCAAGATTAACCTCTCTTCCCGCCAGAGCATTAACAGACAGTGC 1049
 DB AAGTGAATGGGCGCAGTGGCAAGATTAACCTCTCTTCCCGCCAGAGCATTAACAGACAGTGC 960
 QY TGTACATCTTACCTCTGCGACCAAGAGGCGCTCCCGAAGGCTGCTGGATCAGTCACTGA 1109
 DB TGTACATCTTACCTCTGCGACCAAGAGGCGCTCCCGAAGGCTGCTGGATCAGTCACTGA 1020
 QY AGATCTGCAATGCAAGGCGCTTCTATCACTGTGTGTGTCACAGAGAGATGTGATCT 1169
 DB AGATCTGCAATGCAAGGCGCTTCTATCACTGTGTGTGTCACAGAGAGATGTGATCT 1080
 QY ACCCTGCGCTCCCACTTACATCAATGCTGGTTCCTGCTGGGAGTCCGAGGCTGATGG 1229
 DB ACCCTGCGCTCCCACTTACATCAATGCTGGTTCCTGCTGGGAGTCCGAGGCTGATGG 1140
 QY GCATTGGGGGCAAGTGTGTGTAATCCAGATTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1289
 DB GCATTGGGGGCAAGTGTGTGTAATCCAGATTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 QY GCGAGCAGCAGAGGCTGAGGCTGTTCAGTACATTTGGGGAGCTGTGCGGATACCTTTGCA 1349
 DB GCGAGCAGCAGAGGCTGAGGCTGTTCAGTACATTTGGGGAGCTGTGCGGATACCTTTGCA 1260
 QY ACCAGCCCCGAGCAGAGGAGAACTGCGCATTAAGGCTGGAGTGGGCTGAGCGGAG 1409
 DB ACCAGCCCCGAGCAGAGGAGAACTGCGCATTAAGGCTGGAGTGGGCTGAGCGGAG 1320
 QY TGGCCCGAGATCTTGGGAGAGCTTTTGTGCGGCGCTTGGGCGCTTGGAGTGTCTGAGGA 1469
 DB TGGCCCGAGATCTTGGGAGAGCTTTTGTGCGGCGCTTGGGCGCTTGGAGTGTCTGAGGA 1380
 QY TGGCCCGAGATCTTGGGAGAGCTTTTGTGCGGCGCTTGGGCGCTTGGAGTGTCTGAGGA 1529
 DB TGGCCCGAGATCTTGGGAGAGCTTTTGTGCGGCGCTTGGGCGCTTGGAGTGTCTGAGGA 1440
 QY TGGGCGGTGCTTCTGCTTTCAGAGCATATCTTCCCTTCTGCTGATTCCTATGATG 1589
 DB TGGGCGGTGCTTCTGCTTTCAGAGCATATCTTCCCTTCTGCTGATTCCTATGATG 1500
 QY TCAACACAGAGAGCCCAATTCGGGAGCCCGAGGAGAGCTGTATGGCAATCCAGAGG 1649
 DB TCAACACAGAGAGCCCAATTCGGGAGCCCGAGGAGAGCTGTATGGCAATCCAGAGG 1560
 QY AGCCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1709
 DB AGCCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
 QY GGCACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1769
 DB GGCACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
 QY TCAACACTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1829
 DB TCAACACTGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
 QY CTGGAAGACCTTCAAGTGTGAAGGGGAGAAAGTGTGCAACACAGAGTGGGAGAGTGT 1889
 DB CTGGAAGACCTTCAAGTGTGAAGGGGAGAAAGTGTGCAACACAGAGTGGGAGAGTGT 1800
 QY TCGAGGCGCTAGATTTTCTTCAAGAGGTGAAGCTGTGAGTGTGCTGCTGCTGCTGCTGCTG 1949
 DB TCGAGGCGCTAGATTTTCTTCAAGAGGTGAAGCTGTGAGTGTGCTGCTGCTGCTGCTGCTG 1860
 QY AAGGAGGAGCTGGAATGAGAGCTTATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2009
 DB AAGGAGGAGCTGGAATGAGAGCTTATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
 QY AGCTTACACCAACGCTGTGAGAGCTTGTGCACTTATGCGGCGGCGGAGTTCCTCAGGC 2069

Accession	Sequence	Length
Db	1921 AGCTCTACCCACGAGTGTCTGAAACTTGGACCTTATGCCCGGCCCGGATTTCTCAGGC	1950
QY	2070 TCCAGAGATCTTTTGGCCACCAAGAGACCTTTCAAAACAGCAGAAAATTTCGATGGCAATG	2129
Db	1981 TCCAGAGATCTTTTGGCCACCAAGAGACCTTTCAAAACAGCAGAAAATTTCGATGGCAATG	2040
QY	2130 AGGGCTTCGACCCCGACGACCCCTGTCTGATACCCACATGTACGTTCTTGACCCAGGCTGTAGGTG	2189
Db	2041 AGGGCTTCGACCCCGACGACCCCTGTCTGATACCCACATGTACGTTCTTGACCCAGGCTGTAGGTG	2100
QY	2190 CCTACCTTGCCCTTCACAACTGCCCCGGATACAGCGCCCTCTTGACAGAAACCTTGAAATCT	2249
Db	2101 CCTACCTTGCCCTTCACAACTGCCCCGGATACAGCGCCCTCTTGACAGAAACCTTGAAATCT	2160
QY	2250 GAGAACTTCCACACCTTAGAGGACCTGTAGAGAGGAACTGTGTGGGTGGGGGCGGTGCAG	2309
Db	2161 GAGAACTTCCACACCTTAGAGGACCTGTAGAGAGGAACTGTGTGGGTGGGGGCGGTGCAG	2220
QY	2310 GTGTACTGCGGCTGTCAAGGATCTTTTCTATACCAAGACTGGGGTCACTATTTTGAAATA	2369
Db	2221 GTGTACTGCGGCTGTCAAGGATCTTTTCTATACCAAGAACTGGGGTCACTATTTTGAAATA	2280
QY	2370 ATGTGGCTGGAGCTGATCCAGCTGTCTGTGACT	2403
Db	2281 ATGTGGCTGGAGCTGATCCAGCTGTCTGTGACT	2314

RESULT 11			
BD159613			
LOCUS	BD159613	2314 bp	DNA linear PAT 17-JAN-2003
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.		

ACCESSION	BD159613
VERSION	BD159613.1
KEYWORDS	GI:27865371
SOURCE	JP 2002191363-A/14456.
	Homo sapiens (human)

ORGANISM	Homio sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 314)
AUTHORS	Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,...
TITLE	Ishii, S., Sugiyama, T., Wakematsu, A., Nagai, K., and Otsukii, T.
JOURNAL	Primer for synthesizing full-length cDNA and use thereof
	Patent: JP 2002191363-A 14456 09-JUN-2002;

COMMENT
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/14456
PD 09-JUL-2002
PF 28-JUL-2000 JP 20002806990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI TAITO
PI KEIICHI NAGAI, TETSUJI OTSUKI
PI

	C ₆	C ₇	C ₈	C ₉	C ₁₀	C ₁₁	C ₁₂	C ₁₃	C ₁₄	C ₁₅	C ₁₆	C ₁₇	C ₁₈	C ₁₉	C ₂₀	C ₂₁	C ₂₂	C ₂₃	C ₂₄	C ₂₅	C ₂₆	C ₂₇	C ₂₈	C ₂₉	C ₃₀	C ₃₁	C ₃₂	C ₃₃	C ₃₄	C ₃₅	C ₃₆	C ₃₇	C ₃₈	C ₃₉	C ₄₀	C ₄₁	C ₄₂	C ₄₃	C ₄₄	C ₄₅	C ₄₆	C ₄₇	C ₄₈	C ₄₉	C ₅₀	C ₅₁	C ₅₂	C ₅₃	C ₅₄	C ₅₅	C ₅₆	C ₅₇	C ₅₈	C ₅₉	C ₆₀	C ₆₁	C ₆₂	C ₆₃	C ₆₄	C ₆₅	C ₆₆	C ₆₇	C ₆₈	C ₆₉	C ₇₀	C ₇₁	C ₇₂	C ₇₃	C ₇₄	C ₇₅	C ₇₆	C ₇₇	C ₇₈	C ₇₉	C ₈₀	C ₈₁	C ₈₂	C ₈₃	C ₈₄	C ₈₅	C ₈₆	C ₈₇	C ₈₈	C ₈₉	C ₉₀	C ₉₁	C ₉₂	C ₉₃	C ₉₄	C ₉₅	C ₉₆	C ₉₇	C ₉₈	C ₉₉	C ₁₀₀	C ₁₀₁	C ₁₀₂	C ₁₀₃	C ₁₀₄	C ₁₀₅	C ₁₀₆	C ₁₀₇	C ₁₀₈	C ₁₀₉	C ₁₁₀	C ₁₁₁	C ₁₁₂	C ₁₁₃	C ₁₁₄	C ₁₁₅	C ₁₁₆	C ₁₁₇	C ₁₁₈	C ₁₁₉	C ₁₂₀	C ₁₂₁	C ₁₂₂	C ₁₂₃	C ₁₂₄	C ₁₂₅	C ₁₂₆	C ₁₂₇	C ₁₂₈	C ₁₂₉	C ₁₃₀	C ₁₃₁	C ₁₃₂	C ₁₃₃	C ₁₃₄	C ₁₃₅	C ₁₃₆	C ₁₃₇	C ₁₃₈	C ₁₃₉	C ₁₄₀	C ₁₄₁	C ₁₄₂	C ₁₄₃	C ₁₄₄	C ₁₄₅	C ₁₄₆	C ₁₄₇	C ₁₄₈	C ₁₄₉	C ₁₅₀	C ₁₅₁	C ₁₅₂	C ₁₅₃	C ₁₅₄	C ₁₅₅	C ₁₅₆	C ₁₅₇	C ₁₅₈	C ₁₅₉	C ₁₆₀	C ₁₆₁	C ₁₆₂	C ₁₆₃	C ₁₆₄	C ₁₆₅	C ₁₆₆	C ₁₆₇	C ₁₆₈	C ₁₆₉	C ₁₇₀	C ₁₇₁	C ₁₇₂	C ₁₇₃	C ₁₇₄	C ₁₇₅	C ₁₇₆	C ₁₇₇	C ₁₇₈	C ₁₇₉	C ₁₈₀	C ₁₈₁	C ₁₈₂	C ₁₈₃	C ₁₈₄	C ₁₈₅	C ₁₈₆	C ₁₈₇	C ₁₈₈	C ₁₈₉	C ₁₉₀	C ₁₉₁	C ₁₉₂	C ₁₉₃	C ₁₉₄	C ₁₉₅	C ₁₉₆	C ₁₉₇	C ₁₉₈	C ₁₉₉	C ₂₀₀	C ₂₀₁	C ₂₀₂	C ₂₀₃	C ₂₀₄	C ₂₀₅	C ₂₀₆	C ₂₀₇	C ₂₀₈	C ₂₀₉	C ₂₁₀	C ₂₁₁	C ₂₁₂	C ₂₁₃	C ₂₁₄	C ₂₁₅	C ₂₁₆	C ₂₁₇	C ₂₁₈	C ₂₁₉	C ₂₂₀	C ₂₂₁	C ₂₂₂	C ₂₂₃	C ₂₂₄	C ₂₂₅	C ₂₂₆	C ₂₂₇	C ₂₂₈	C ₂₂₉	C ₂₃₀	C ₂₃₁	C ₂₃₂	C ₂₃₃	C ₂₃₄	C ₂₃₅	C ₂₃₆	C ₂₃₇	C ₂₃₈	C ₂₃₉	C ₂₄₀	C ₂₄₁	C ₂₄₂	C ₂₄₃	C ₂₄₄	C ₂₄₅	C ₂₄₆	C ₂₄₇	C ₂₄₈	C ₂₄₉	C ₂₅₀	C ₂₅₁	C ₂₅₂	C ₂₅₃	C ₂₅₄	C ₂₅₅	C ₂₅₆	C ₂₅₇	C ₂₅₈	C ₂₅₉	C ₂₆₀	C ₂₆₁	C ₂₆₂	C ₂₆₃	C ₂₆₄	C ₂₆₅	C ₂₆₆	C ₂₆₇	C ₂₆₈	C ₂₆₉	C ₂₇₀	C ₂₇₁	C ₂₇₂	C ₂₇₃	C ₂₇₄	C ₂₇₅	C ₂₇₆	C ₂₇₇	C ₂₇₈	C ₂₇₉	C ₂₈₀	C ₂₈₁	C ₂₈₂	C ₂₈₃	C ₂₈₄	C ₂₈₅	C ₂₈₆	C ₂₈₇	C ₂₈₈	C ₂₈₉	C ₂₉₀	C ₂₉₁	C ₂₉₂	C ₂₉₃	C ₂₉₄	C ₂₉₅	C ₂₉₆	C ₂₉₇	C ₂₉₈	C ₂₉₉	C ₃₀₀	C ₃₀₁	C ₃₀₂	C ₃₀₃	C ₃₀₄	C ₃₀₅	C ₃₀₆	C ₃₀₇	C ₃₀₈	C ₃₀₉	C ₃₁₀	C ₃₁₁	C ₃₁₂	C ₃₁₃	C ₃₁₄	C ₃₁₅	C ₃₁₆	C ₃₁₇	C ₃₁₈	C ₃₁₉	C ₃₂₀	C ₃₂₁	C ₃₂₂	C ₃₂₃	C ₃₂₄	C ₃₂₅	C ₃₂₆	C ₃₂₇
--	----------------	----------------	----------------	----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------	------------------

```

FEATURES
source
location/Qualifiers
1..2314
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

ORIGIN	Query Match	Score	DB	Length
Best Local Similarity	99.9%	Pred. No. 0		
Matches 212; Conservative	0	Mismatches 2	Indels 0	Gaps 0
90	GGAGAGAAAGTCTCAGCTAGACGAGGGGCGCTTAGTTTGGAAAGGAGGATCAGGGA	149		

D	b	I	GGAAGAGAAGTCTCAGACTAGAACGACGGGCCATTGTTTCGAAAGSAGAGATCAGGA	60
O	y	150	TGTTTGCGAGCGCTGGAAACCAGCGGTGCCGATAGAGAAAGCGGGCTCCATGGCTGCC	209
D	b	61	TGTTTGCGAGCGCTGGAAACCAGCGGTGCCGATAGAGAAAGCGGGCTCCATGGCTGCC	120
O	y	210	TTCCGCGCTGCCCCCTGCTGCTGTGTGTGCATCAACCGTGCCTGCTGCTAAGTCAACCTCTGAC	269
D	b	121	TTCTGCTGCTGCCCCCTGCTGCTGTGTGTGCATCAACCGTGCCTGCTGCTAAGTCAACCTCTGAC	180
O	y	270	CGCAGTTGCGCTGAGCTTCCGGCGGACTTGGCCCTTTCGGGTGCGAGCTCTGTGTCAAAA	329
D	b	181	CGAGTTTGCGCTGAGCTTCCGGCGGACTTGGCCCTTTCGGGTGCGAGCTCTGTGTCAAAA	240
O	y	330	GAGGCTTTTCCAGCTTCGGGCTTGGCCCGCGGCTGGCCCGCGAACCGGAAGSTCCAGAGGGG	389
D	b	241	GAGGCTTTTCCAGCTTCGGGCTTGGCCCGCGGCTGGCCCGAACCGGAAGSTCCAGAGGGG	300
O	y	390	GCTCAGCGCTTAGCGCTCTCCCGGAACTGCGCCACAGCAGCGCGCCGCGACACTTTTC	449
D	b	301	GCTCAGCGCTTAGCGCTCTCCCGGAACTGCGCCACAGCAGCGCGCCGCGACACTTTTC	360
O	y	450	TCATTCACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGGTAAACAGGCTG	509
D	b	361	TCATTCACGGCTCGCGGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGGTAAACAGGCTG	420
O	y	510	CAGCGGCTTTCCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGCGCGACGCGCG	569
D	b	421	CACCGGCTTTCCTACGTGCGCTAGGCTGGGACTGGGGACCCGACGGCGCGCGACGCGCG	480
O	y	570	AGGGAGCGCTGAGAAAGCGAGCGCGGACGCGCGGAGCCGGAATGCAAGCGCGGAA	629
D	b	481	AGGGAGCGCTGAGAAAGCGAGCGCGGACGCGCGGAGCCGGAATGCAAGCGCGGAA	540
O	y	630	GCGGCGCGGAGTTTGGCGGAGGGGACGGTGGCCGACAGGTGGAGAGACCCCGGCGCTC	689
D	b	541	GCGGCGCGGAGTTTGGCGGAGGGGACGGTGGCCGACAGGTGGAGAGACCCCGGCGCTC	600
O	y	690	TGTACCTGGAGCAACTGTGGCGTGCTCTCTCCCGCTGCGGCCCAAGTTCGTGGCTCT	749
D	b	601	TGTACCTGGAGCAACTGTGGCGTGCTCTCTCTCCCGCTGCGGCCCAAGTTCGTGGCTCT	660
O	y	750	GATTGGGCTGGCCAAGGCGCGGCTTGCGCACTGCTTTGTGCCACCGCCCTTGCGCGCGG	809
D	b	661	GATTGGGCTGGCCAAGGCGCGGCTTGCGCACTGCTTTGTGCCACCGCCCTTGCGCGCGG	720
O	y	810	GGCCCGTGGCACTGGCTTCGCGAGCTGCGGCGCGCGCGCGCTGGTGTGCGGCGCGAGT	869
D	b	721	GGCCCGTGGCACTGGCTTCGCGAGCTGCGGCGCGCGCGCGCTGGTGTGCGGCGCGAGT	780
O	y	870	TTCTGGAATCCCTGSAAGCGGACCTGCCCAGCTTAAGAGCATGGGGCTTCCACTGTGGG	929
D	b	781	TTCTGGAATCCCTGSAAGCGGACCTGCCCAGCTTAAGAGCATGGGGCTTCCACTGTGGG	840
O	y	930	CTGCAAGGCCCAAGMAACCACTCTGTGTAATTAAGATTTTCTGGCTMAAGTGTCCGCTG	989
D	b	841	CTGCAAGGCCCAAGMAACCACTCTGTGTAATTAAGATTTTCTGGCTMAAGTGTCCGCTG	900
O	y	990	AAGTGAATGAGCACAGTGGCAGATACTCTTTCCCGCCAGAGCATTAACAGACAGTGGC	1044
D	b	901	AAGTGAATGAGCACAGTGGCAGATACTCTTTCCCGCCAGAGCATTAACAGACAGTGGC	960
O	y	1050	TGTACATTTCACTTGGCACCAAGGGGCTTCCCAAGGCTGTGTGATCACTCTGA	1104
D	b	961	TGTACATTTCACTTGGCACCAAGGGGCTTCCCAAGGCTGTGTGATCACTCTGA	1020
O	y	1110	AGATTCGGAATGGCAGGGGCTTTCATACAGTGTGTGTGTGCACACAGGAAGTGAATCT	1165
D	b	1021	AGATTCGGAATGGCAGGGGCTTTCATACAGTGTGTGTGTGCACACAGGAAGTGAATCT	1080
O	y	1170	ACCTGCGCCTCCACTACCAATGTGCGGTTCCCTGCTGAGGACTGTGGGCTGCAATGG	1222
D	b	1081	ACCTGCGCCTCCACTACCAATGTGCGGTTCCCTGCTGAGGACTGTGGGCTGCAATGG	1144

QY 1230 GCATTGGGGCCACAGTGTGCTGAATCCAAATTCCTGGCTGTGCAATTTTGGGAAAGATT 1289
DB 1141 GGAATGGGGCCACAGTGTGCTGAATCCAAATTCCTGGCTGTGCAATTTTGGGAAAGATT 1200
QY 1290 GCAGAGAGACAGGGGAGAGGTTCCAGTACATTTGGGGAGCTGAGCCGATACCTTCA 1349
DB 1201 GCCAGACACAGGGTGAAGGTTCCAGTACATTTGGGGAGCTGAGCCGATACCTTCA 1260
QY 1350 ACCAGCCCCGAGCAGAGGAGAGAGCTGAGCCATAGGTCCGCTGGCAGTGGGACAGGGGC 1409
DB 1261 ACCAGCCCCGAGCAGAGGAGAGAGCTGAGCCATAGGTCCGCTGGCAGTGGGACAGGGGC 1340
QY 1410 TGGGCCAATACCTTGGGAGGCTTTTGGGCGGCTTGGGGCCCTGAGAGGTGTGAGGA 1469
DB 1321 TGGGCCAATACCTTGGGAGGCTTTTGGGCGGCTTGGGGCCCTGAGAGGTGTGAGGA 1380
QY 1470 CATATGACTGACAGAGGAGCAGTGGCCCATCATCACTACACAGAGAGCGGGGCGCTG 1529
DB 1381 CATATGACTGACAGAGGAGCAGTGGCCCATCATCACTACACAGAGAGCGGGGCGCTG 1440
QY 1530 TGGGGGTGCTCTTCTGGCTTTCAAGCATATCTTCCCTCTCTGATTTGGCTATGATG 1589
DB 1441 TGGGGGTGCTCTTCTGGCTTTCAAGCATATCTTCCCTCTCTGATTTGGCTATGATG 1500
QY 1590 TCACCAAGAGAGAGCAGTGGGAGCCCGCAGGGGACCTGATGGCCACATCTCCAGGTG 1649
DB 1501 TCACCAAGAGAGAGCAGTGGGAGCCCGCAGGGGACCTGATGGCCACATCTCCAGGTG 1560
QY 1650 AGCCAGGCTGCTGTGGTGGCCCGGTAAGCAGCAGTCCCATCTCTGAGCTATGCTGAG 1709
DB 1561 AGCCAGGCTGCTGTGGTGGCCCGGTAAGCAGCAGTCCCATCTCTGAGCTATGCTGAG 1620
QY 1710 GGGCAGAGCTGGCCCGCAGGGGAAAGTTGCTAAAGATGCTTCCGGCTGGGGATGTTTCT 1769
DB 1621 GGGCAGAGCTGGCCCGCAGGGGAAAGTTGCTAAAGATGCTTCCGGCTGGGGATGTTTCT 1680
QY 1770 TCACCACTGGGAGACCTGCTGTGTCTGATGACCAAGTTTCTCCGCTTCATGATGTA 1829
DB 1681 TCACCACTGGGAGACCTGCTGTGTCTGATGACCAAGTTTCTCCGCTTCATGATGTA 1740
QY 1830 CTGGAGACACCTTCAAGTGGAGAGGGGAGAAATGTGCCACCAACGAGGTGGCAGAGGCT 1889
DB 1741 CTGGAGACACCTTCAAGTGGAGAGGGGAGAAATGTGCCACCAACGAGGTGGCAGAGGCT 1800
QY 1890 TCGAGGCTCTAGATTTTCTTCAAGAGTGAACGTCTATGATGATCTGTGCCAGGGCATG 1949
DB 1801 TCGAGGCTCTAGATTTTCTTCAAGAGTGAACGTCTATGATGATCTGTGCCAGGGCATG 1860
QY 1950 AAGGCAAGGCTGGAATGGCAAGCCCTAGTTTGTGCTCCCGCCAGCTTTGACCTTATGC 2009
DB 1861 AAGGCAAGGCTGGAATGGCAAGCCCTAGTTTGTGCTCCCGCCAGCTTTGACCTTATGC 1920
QY 2010 AACTCTACACCAAGCTGTGTGAGAACTTGCACCTTATGCGCGGCCGATTCCTCAAGGC 2069
DB 1921 AACTCTACACCAAGCTGTGTGAGAACTTGCACCTTATGCGCGGCCGATTCCTCAAGGC 1980
QY 2070 TCCAGAGGCTTTTGGCCACACAGAGACCTTCAACAGAGAAATTTGATGGCAATG 2129
DB 1981 TCCAGAGGCTTTTGGCCACACAGAGACCTTCAACAGAGAAATTTGATGGCAATG 2040
QY 2130 AAGGCTTGGACCCAGCAACCTGTCTGACCCACCTGATGTTTGGACAGGCTGTAGGTG 2189
DB 2041 AAGGCTTGGACCCAGCAACCTGTCTGACCCACCTGATGTTTGGACAGGCTGTAGGTG 2100
QY 2190 CCTACTGTGCTTCAACATGTGCGCGGTACAGGCGCTCTGAGCAAGAACTTGTGAATCT 2249
DB 2101 CCTACTGTGCTTCAACATGTGCGCGGTACAGGCGCTCTGAGCAAGAACTTGTGAATCT 2160
QY 2250 GAGAACTTCCACACTGAGAGGACCTGAGAGAGGAACTCTGTGGGTGGGGGCGCTGTAG 2309
DB 2161 GAGAACTTCCACACTGAGAGGACCTGAGAGAGGAACTCTGTGGGTGGGGGCGCTGTAG 2220

QY 2310 GTGTACTGGGCTGTCAAGGATCTTTTCTATACCAGATCGCGTCACTATTTTGTATAA 2369
DB 2221 GTGTACTGGGCTGTCAAGGATCTTTTCTATACCAGATCGCGTCACTATTTTGTATAA 2280
QY 2370 ATGAGGCTGAGCTGATCCAGTGTCTGACCT 2403
DB 2281 ATGAGGCTGAGCTGATCCAGTGTCTGACCT 2314
RESULT 12
AK027499
LOCUS
DEFINITION
Homo sapiens cDNA FLJ14593 fis, clone NTBRM4002073, moderately similar to Mus musculus fatty acid transport protein 3 mRNA.
ACCESSION
AK027499.1 GI:14042218
VERSION
AK027499.1
KEYWORDS
Oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
1 Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Magatsuma, M., Hosoliri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, X., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Makamatsu, A., Nakamura, Y., Nagahara, K., Masuo, Y., Ninomiya, K. and Iwayanagi, T.
TITLE
NEDO human cDNA sequencing project
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2314)
AUTHORS
Isogai, T. and Otsuki, T.
TITLE
Direct Submission
JOURNAL
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kizatsuru Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES
Location/Qualifiers
source
1. 2314
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM4002073"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RM4"
/note="Cloning vector: pUC19FL3-mRNA from uninduced NT2 neuronal precursor cells."
CDS
60..2162
/note="unlabeled protein product"
/codon_start=1
/protein_id="BAB5156.1"
/db_xref="GI:14042218"
/translation="MFASGMNQTVIEBGSVAALLLLPLLLPLLLKXHLWPEOLWLPADLAFVRAICCKRAARALAAADPEPGGCSLARLAEILAOQRAATFLIHGSRPVSSEARERENRARAFRLALGMDWGDGDSGSGASGEGERAAPGDAAGSABRPGCGDGAARGGAAAPLSPGATVALLPAGSEPLIMWGLKALRLAFVFTALRGPPLHCHRCSCARALVLAPEFLSLSEBDLPALAMGLHMAAPGTHRPAGISLLAEVSAEVDGPVGLSSPQITDTCLYIFSGITGLPKARISHLKILOCQGFYOLGVHOEDVLYALPLYHMSGSLGIVGCMGATVLLSKRSAGQFWDCCQHRVTVQYIGELCYLVQPPSKAERGHVLAAGSGIAPRTWFRFVRRPDLQVLETFYGLTENVA TINYQORGAVERASMLYKHIPFSLIRDVTTGPIIDPOGCVATPGEGRGLVAVVQOSPFGLYAGVGPILAOGSKLLKDFRGGDVEFNTGDLVCDQGFLEFHEBDTPTFRWGENVATREVAAREALDPLQDVNVVCGTVPGHGRGAMALVLRPHALDLMQLYTHSENIPYARPRFLRQDSLATITETFEQOKVRANEGFDSTLSLPLVLDQAVAYLPLTARYSALLAGNLRIT"

QY 2190 CCTACTGCGCTGCAACTGCGGCTAGAGGCGCTCTGCGAGAAACCTTGAATCT 2249
 DB 2101 CCTACTGCGCTGCAACTGCGGCTAGAGGCGCTCTGCGAGAAACCTTGAATCT 2160
 QY 2250 GAGAACTTCACACCTGAGGACCTGAGAGAGAACTCTGTGGGGTGGGGCGCTTGCAG 2309
 DB 2161 GAGAACTTCACACCTGAGGACCTGAGAGAGAACTCTGTGGGGTGGGGCGCTTGCAG 2220
 QY 2310 GTGTACTGGGCTGTGAGGAGATCTTTCTATACCAACCTGGCGTCACTATTTTGAATTA 2369
 DB 2221 GTGTACTGGGCTGTGAGGAGATCTTTCTATACCAACCTGGCGTCACTATTTTGAATTA 2280
 QY 2370 ATGTGGCTGAGCTGATCCAGCTGTCTCTGACCT 2403
 DB 2281 ATGTGGCTGAGCTGATCCAGCTGTCTCTGACCT 2314

RESULT 13
 BC029792 2333 bp mRNA linear PRI 07-OCT-2003
 LOCUS Homo sapiens solute carrier family 27 (fatty acid transporter), member 3, mRNA (cDNA clone MGC:35483 IMAGE:5195359), complete cds.
 DEFINITION BC029792
 ACCESSION BC029792.1 GI:20988247
 VERSION MGC.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 2333)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stajdler, E., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Frange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McMan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, X., S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, R.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitt, J., Myers, R.M., Buetow, K.H., Schein, J.E., Jones, S.J., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL 22388257
 MEDLINE 12477932
 PubMed 2 (bases 1 to 2333)
 REFERENCE Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILN)
 DNA Sequencing Center: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hui, X., S.W., Loulseged, H., Kowis, C.R., Sneed, A.U., Martin, R.G., Muzny, D.M., Nataravali,

A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN at: <http://image.llnl.gov>
 Series: IRAX Plate: 50 Row: n Column: 23.
 Location/Qualifiers
 1. 2333
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:35483 IMAGE:5195359"
 /tissue_type="Brain, adult, 6 pooled whole brains"
 /clone_lib="NIH MGC_114"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORE"
 1. 2333
 /gene="SLC27A3"
 /note="synonyms: FATP3, MGC4365"
 /db_xref="LOCUSID:11000"
 /db_xref="MIM:604193"
 24. 2126
 /codon_start=1
 /product="SLC27A3 protein"
 /protein_id="AA29792.1"
 /db_xref="GI:20988248"
 /db_xref="LOCUSID:11000"
 /translation="MPSGQNTVPPIEAGSMAALLPLLLPLLLKLLHLPQLR WLPADIAFVRLCKKRALRALAAADPGPGGSGSLAWRLLELAQOAAHFLI HGRSPRSYSEAEBSNRARAFRALGMDGSDGSDGSGSEGRAPAGDAAAL SGAFPGAGGAAGGAAPLSPGATVALLPAGPEPLWFLWFLKAGARTAFVPL RRGPLHCRSCGAPALVLAPELSEEDLPLAMGCHLMAAGPCTHAPDISDLA EYSAEDVGPVPLSPDITDCLIFPSTGTGPKARISHLILQCGYQLQGV HOEDVLYLPLPYHMSGLIGVCGIGATVLLSKSASGFMEDCOQRTVFOYI GELCRVLAQVPSKRGKRLVAAGSGRPPTWRPVRFGLVLEHYGATLSENVAV TIVYTORQAVRASLWLYKIFPFLIRYDVTGPIRDPQCHVATSGEGLVAP VSOQSPFLVAGPGLAOKLKDVERPDVFPNFGDLVCDQGFRLHEDTGTFR WGENVATVEVALEFEALDPLEVAVVGVTVGHRGAGMALVLRPPALDMLQYT HUSENIPPRARRRLRLQSLATTFTRFKQKRVMANEGFDEPLDPLVLDQAVAY LPLTRKISALAGNLR"
 579. 2015
 /note="Caic; Region: Acyl-CoA synthetases
 (AMP-forming)/AMP-acid ligases II [lipid metabolism / Secondary metabolites biosynthesis, transport, and catabolism]"
 /db_xref="CCD:COG0318"

ORIGIN
 Query Match 94.8%; Score 2280; DB 9; Length 2333;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 GTTTGGAGAGAGATGAGGATGTTTCCAGAGCGCTGGAAACGAGCGTGGCGATG 185
 DB 1 GTTTGGAGAGAGATGAGGATGTTTCCAGAGCGCTGGAAACGAGCGTGGCGATG 60
 QY 186 AGAAGCGGCTCATGAGCTGCGCTCTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTG 245
 DB 61 AGAAGCGGCTCATGAGCTGCGCTCTGCTGCTGCGCTGCTGCTGCTGCTGCTGCTG 120
 QY 246 TGTCTGAACTACACCTTGGCGCGAGTGTGCGCTGCTGCGCGAGCTTGGCGCTTTG 305
 DB 121 TGTCTGAACTACACCTTGGCGCGAGTGTGCGCTGCTGCGCGAGCTTGGCGCTTTG 180
 QY 306 CGGTGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 365
 DB 181 CGGTGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 QY 366 CGAAGCGGAGTCCGAGAGGGGCTGAGAGCTTGGCGCTGCGCTGCGGAACTGGCGCC 425
 DB 241 CGAAGCGGAGTCCGAGAGGGGCTGAGAGCTTGGCGCTGCGCTGCGGAACTGGCGCC 300
 QY 426 AGAAGCGGCGCGGACACCTTTCATTCAGAGCTGCGCGGCGCTTACTACTGAGAG 485

Db 301 AGCAGCGCGCGCAGACCTTTCTCATTCAGCGCTCGCGCGCTTTAGCTACAGAGG 360
Qy 486 CGAGCGCGAGAGTAACAGGCGCTGACGCGCTTCTCACTGCGCTAGGCTGGAGTGGG 545
Db 361 CGAGCGCGAGAGTAACAGGCGCTGACGCGCTTCTCACTGCGCTAGGCTGGAGTGGG 420
Qy 546 GACCCGACGCGCGCAGCAGCGCGAGCGCTGAGAAAGCGAGCGGCGAGCGCGG 605
Db 421 GACCCGACGCGCGCAGCAGCGCGAGCGCTGAGAAAGCGAGCGGCGAGCGCGG 480
Qy 606 GAGCGCGAGATGACAGCGCGCGAGCGCGAGCGCTGAGAAAGCGAGCGGCGAGCGG 665
Db 481 GAGCGCGAGATGACAGCGCGCGAGCGCGAGCGCTGAGAAAGCGAGCGGCGAGCGG 540
Qy 666 GAGCGCGAGATGACAGCGCGCGAGCGCGAGCGCTGAGAAAGCGAGCGGCGAGCGG 725
Db 541 GAGCGCGAGATGACAGCGCGCGAGCGCGAGCGCTGAGAAAGCGAGCGGCGAGCGG 600
Qy 726 CTGCGCGCAGAGTTTCTGCGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 785
Db 601 CTGCGCGCAGAGTTTCTGCGCTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 660
Qy 786 TTGTCGCCACGCGCTGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCG 845
Db 661 TTGTCGCCACGCGCTGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCG 720
Qy 846 GCGCGCTGCTGCTGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCG 905
Db 721 GCGCGCTGCTGCTGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCG 780
Qy 906 GAGCGATGCGCGCTGCGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCT 965
Db 781 GAGCGATGCGCGCTGCGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCT 840
Qy 966 ATTGCTGCGCTGCGCTGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCT 1025
Db 841 ATTGCTGCGCTGCGCTGCGCGCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCT 900
Qy 1026 CCCAAGCATACACACACGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1085
Db 901 CCCAAGCATACACACACGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 960
Qy 1086 AGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1145
Db 961 AGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1020
Qy 1146 GTGTCCACCGAGAAATGATGATCTACCTGCGCTTCCACTCTACCAATGTCGGTTC 1205
Db 1021 GTGTCCACCGAGAAATGATGATCTACCTGCGCTTCCACTCTACCAATGTCGGTTC 1080
Qy 1206 TCGTGGGACGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1265
Db 1081 TCGTGGGACGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1140
Qy 1266 CGCGCTGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1325
Db 1141 CGCGCTGCTGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1200
Qy 1326 GGGAGCTGCGCGATACCTGTCAACCAAGCGCGCGAGAAAGCGAGAAAGCGAGT 1385
Db 1201 GGGAGCTGCGCGATACCTGTCAACCAAGCGCGCGAGAAAGCGAGAAAGCGAGT 1260
Qy 1386 TCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1445
Db 1261 TCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1320
Qy 1446 TCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1505
Db 1321 TCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1380
Qy 1506 ACTACACAGACAGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1565
Db 1381 ACTACACAGACAGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1440

Qy 1566 CCTCTCTTGATTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1625
Db 1441 CCTCTCTTGATTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Qy 1626 ACTGATGCGCAATCTCCAGGTGACCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1685
Db 1501 ACTGATGCGCAATCTCCAGGTGACCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Qy 1686 CCCATTCTCGGCTATGCTGCGCGCGCAGAGCTGCGCGCAGAGGAAATTGCTAAAGATG 1745
Db 1561 CCCATTCTCGGCTATGCTGCGCGCGCAGAGCTGCGCGCAGAGGAAATTGCTAAAGATG 1620
Qy 1746 TCTTCCGCGCTGGGAGATTTTCTTCAACATGCGGGAATGCTGCTGCTGCTGCTGCT 1805
Db 1621 TCTTCCGCGCTGGGAGATTTTCTTCAACATGCGGGAATGCTGCTGCTGCTGCTGCT 1680
Qy 1806 GTTTCTCGGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1865
Db 1681 GTTTCTCGGCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
Qy 1866 CCAACCGAGTGGCAGAGGCTTTCAGAGCGCTGAGATTTTCTTCAAGAGTGAACGTCT 1925
Db 1741 CCAACCGAGTGGCAGAGGCTTTCAGAGCGCTGAGATTTTCTTCAAGAGTGAACGTCT 1800
Qy 1926 ATGAGTCACTGTCAGAGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1985
Db 1801 ATGAGTCACTGTCAGAGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
Qy 1986 CCCCCAGCTTTGAGACCTTATGACAGCTCTACACCCAGCTGTGAGAACTTTCACCTT 2045
Db 1861 CCCCCAGCTTTGAGACCTTATGACAGCTCTACACCCAGCTGTGAGAACTTTCACCTT 1920
Qy 2046 ATGCGCGCGCGCGATTTCTGAGGCTCCAGAGCTTTTGGCCACCAAGAGACCTTAAAC 2105
Db 1921 ATGCGCGCGCGCGATTTCTGAGGCTCCAGAGCTTTTGGCCACCAAGAGACCTTAAAC 1980
Qy 2106 AGCAGAAAGTTCGATGCGCAATGAGGAGGCTGAGCCCGACCCCTGTGAGCCACTGT 2165
Db 1981 AGCAGAAAGTTCGATGCGCAATGAGGAGGCTGAGCCCGACCCCTGTGAGCCACTGT 2040
Qy 2166 AGCTTTCGACAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2225
Db 2041 AGCTTTCGACAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Qy 2226 TCGTGGAGAAACCTTTCGAACTGAGAACTTCCACACCTGAGAGCACTGAGAGAGAAC 2285
Db 2101 TCGTGGAGAAACCTTTCGAACTGAGAACTTCCACACCTGAGAGCACTGAGAGAGAAC 2160
Qy 2286 TCGTGGAGTGGGCGGCTGAGAGTGAATGAGGCTGTCAGAGGATCTTTTCTATACAGA 2345
Db 2161 TCGTGGAGTGGGCGGCTGAGAGTGAATGAGGCTGTCAGAGGATCTTTTCTATACAGA 2220
Qy 2346 ACTGCGTCACTATTTTGTATTAATGCTGCTGAGAGTGAATGAGGCTGCTGAGAGT 2405
Db 2221 ACTGCGTCACTATTTTGTATTAATGCTGCTGAGAGTGAATGAGGCTGCTGAGAGT 2280

RESULT 14
AR168916
LOCUS AR168916 2166 bp DNA linear
DEFINITION Sequence 101 from patent US 6288213.
ACCESSION AR168916
VERSION AR168916.1 GI:17905086
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2166)
AUTHORS Stahl,A., Hirsch,D.J., Lodish,H.F., Glemo,R.E. and Tartaglia,L.A.
TITLE Fatty acid transport proteins
JOURNAL Patent: US 6288213-A 101 11-SEP-2001;
FEATURES Location/Qualifiers

Db 2054 AGGGCTGAGTGCCTACCTGCGCCCTCCACACTGCGCGGTACAGCGCCCTCTGCGAGAA 2113
Qy 2238 ACCTTGAAATCTGAGAACTTCCACACCTGAGAGCACTGAGAGAGAACTGT 2290
Db 2114 ACCTTGAAATCTGAGAACTTCCACACCTGAGAGCACTGAGAGAGAACTGT 2166

RESULT 15
AR193703 2166 bp DNA linear PAT 20-APR-2002
LOCUS AR193703
DEFINITION Sequence 101 from patent US 6348321.
ACCESSION AR193703
VERSION AR193703.1 GI:20240295
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2166)
AUTHORS Stahl, A., Hirsch, D.J., Lodish, H.F., Gimeno, R.E. and Tarragila, L.A.
TITLES Methods of identifying agents inhibiting fatty acid transport proteins
JOURNAL Patent: US 6348321-A 101 19-FEB-2002;
FEATURES Location/Qualifiers
source 1..2166
/organism="Unknown";
/mol_type="unassigned DNA"

ORIGIN
Query Match 88.1%; Score 2119.4; DB 6; Length 2166;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2146; Conservative 0; Mismatches 1; Indels 6; Gaps 2;

Qy 738 TTCTGTGCTGTGCTTCCGAGCTGCGCAAGCCGCGCTGCGCACTGCTTTGTCACAG 797
Db 614 TTCTGTGCTGTGCTTCCGAGCTGCGCAAGCCGCGCTGCGCACTGCTTTGTCACAG 673
Qy 798 CCTGCGCGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
Db 674 CCTGCGCGGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
Qy 858 TGGCGCCAGAGTTCTGAGTCCCTGAGCCGAGCCCTGCGCCCTGAGAGCCATGGGCG 917
Db 734 TGGCGCCAGAGTTCTGAGTCCCTGAGCCGAGCCCTGCGCCCTGAGAGCCATGGGCG 793
Qy 918 TCCACCTGTGGGTGTGAGGCGCCAGAGAGCCACCGCTGAGAAATTAGAGATTGCTGGCTG 977
Db 794 TCCACCTGTGGGTGTGAGGCGCCAGAGAGCCACCGCTGAGAAATTAGAGATTGCTGGCTG 853
Qy 978 AAGTGTCCGCTGAGTGTGAGTGGGCGAGTGCAGAGATACCTCTTCCGCGCCAGAGATTA 1037
Db 854 AAGTGTCCGCTGAGTGTGAGTGGGCGAGTGCAGAGATACCTCTTCCGCGCCAGAGATTA 913
Qy 1038 CAGACAGTGCCTGTGATCTTCACTGCTGAGACCAAGGCGCTCCGCGAGGCTGCTGCGA 1097
Db 914 CAGACAGTGCCTGTGATCTTCACTGCTGAGACCAAGGCGCTCCGCGAGGCTGCTGCGA 973
Qy 1098 TCAGTATCTGAGATCTGCAATGCGCAATGCGCGCTTCTATGAGCTGTGTGTCACAG 1157
Db 974 TCAGTATCTGAGATCTGCAATGCGCAATGCGCGCTTCTATGAGCTGTGTGTCACAG 1033
Qy 1158 AAGATGTATCTAATCTGCGCCCTCCCACTTACCAATGTCGGTTCCTGCTGGGATG 1217
Db 1034 AAGATGTATCTAATCTGCGCCCTCCCACTTACCAATGTCGGTTCCTGCTGGGATG 1093
Qy 1218 TGGGCTGATATGAGGATTTGGGCGCCAGAGTGTGTGTAATTCAGATTCTGCGCTGCTGAT 1277
Db 1094 TGGGCTGATATGAGGATTTGGGCGCCAGAGTGTGTGTAATTCAGATTCTGCGCTGCTGAT 1153
Qy 1278 TCTGGGAAGATTGCCAGACAGACAGAGGTACGGTGTTCAGATACATTGGGAGCTGTGCC 1337
Db 1154 TCTGGGAAGATTGCCAGACAGACAGAGGTACGGTGTTCAGATACATTGGGAGCTGTGCC 1213
Qy 1338 GATACCTGTGCAACAGAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1397
Db 1214 GATACCTGTGCAACAGAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1273
Qy 1398 TGGGAGACGCGGCTGCGCCAGATACCTGAGAGCGTTTGTGCGCGCTTGGGCGCCCTGC 1457
Db 1274 TGGGAGACGCGGCTGCGCCAGATACCTGAGAGCGTTTGTGCGCGCTTGGGCGCCCTGC 1333
Qy 1458 AGGTGCTGAGACATATGATCTGACAGAGGCAAGTGGCCAGCATCACTACACAGAGAC 1517
Db 1334 AGGTGCTGAGACATATGATCTGACAGAGGCAAGTGGCCAGCATCACTACACAGAGAC 1393
Qy 1518 AGCGGGGCGCTGTGGGCGGCTTCTGCTGCTTACAGAGATATCTTCCCTTCTGTA 1577
Db 1394 AGCGGGGCGCTGTGGGCGGCTTCTGCTGCTTACAGAGATATCTTCCCTTCTGTA 1453
Qy 1578 TTGCTATGATGTACCAACAGAGAGCAATTTGGGAGCCCGCGAGGCACTGTATGGCCA 1637
Db 1454 TTGCTATGATGTACCAACAGAGAGCAATTTGGGAGCCCGCGAGGCACTGTATGGCCA 1513
Qy 1638 CATCTCAGGTGAGCCAGAGGCTGCTGTGGGCGCGGTAAGCCAGAGTCCCATTTCTGGA 1697
Db 1514 CATCTCAGGTGAGCCAGAGGCTGCTGTGGGCGCGGTAAGCCAGAGTCCCATTTCTGGA 1573
Qy 1698 GCTATGCTGGGCGGAGAGCTGAGCCCAAGGGAAGTTGCTAAAGATGTCTTCGCGCTTG 1757
Db 1574 GCTATGCTGGGCGGAGAGCTGAGCCCAAGGGAAGTTGCTAAAGATGTCTTCGCGCTTG 1633
Qy 1758 GGGATGTTTCTTCAACATGAGGAGACCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1817
Db 1634 GGGATGTTTCTTCAACATGAGGAGACCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1693
Qy 1818 TCATGATCTGATCTGAGACACCTTCAAGTGTGAGAGGAGAGATGTGGCCACACAGAG 1877

```
Db 1694 TCCATATCGTACTGAGACACCTTCAAGTGGAGGGGAGAAATGTGGCCACACCGAGG 1753
QY 1878 TGGCAGAGGTCTTTCGAGGCCCTAGATTTTCTTCAAGAGGTGAACGTCTATGGAGTCACTG 1937
Db 1754 TGGCAGAGGTCTTTCGAGGCCCTAGATTTTCTTCAAGAGGTGAACGTCTATGGAGTCACTG 1813
QY 1938 TGGCAGAGGTGAAGGCAAGGCTGGAAATGGCAGCCCTAGTTCTGGTCCCCCCCAAGCTT 1997
Db 1814 TGGCAGAGGTGAAGGCAAGGCTGGAAATGGCAGCCCTAGTTCTGGTCCCCCCCAAGCTT 1873
QY 1998 TGGACCTTATGAGCTCTACACCCACGTGTGAGAACTTGGCACCTTATGCCCGGCCCC 2057
Db 1874 TGGACCTTATGAGCTCTACACCCACGTGTGAGAACTTGGCACCTTATGCCCGGCCCC 1933
QY 2058 GATTCTCTAGGCTCCGAGAGTCTTTGGCCACCAAGAGACTTCAAAACAGCAGAAAGTTC 2117
Db 1934 GATTCTCTAGGCTCCGAGAGTCTTTGGCCACCAAGAGACTTCAAAACAGCAGAAAGTTC 1993
QY 2118 GGAATGCAATATGAGGCTTCGACCCGAGCACCCTGTGACCCACTGTAGCTTGGACC 2177
Db 1994 GGAATGCAATATGAGGCTTCGACCCGAGCACCCTGTGACCCACTGTAGCTTGGACC 2053
QY 2178 AGGCTGTAGGTGCTTACTGCCCCCTCACAATGCCCCGTTACAGGCCCTCTGGCAGAA 2237
Db 2054 AGGCTGTAGGTGCTTACTGCCCCCTCACAATGCCCCGTTACAGGCCCTCTGGCAGAA 2113
QY 2238 ACCTTGAACTTGAGAACTTCCACAACCTGAGGCACTGAGAGAGGAACCTGT 2290
Db 2114 ACCTTGAACTTGAGAACTTCCACAACCTGAGGCACTGAGAGAGGAACCTGT 2166
```

Search completed: July 3, 2004, 02:03:09
Job time : 9307 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 22:59:20 ; Search time 6104 Seconds
(without alignments)
11765.817 Million cell updates/sec

Title: US-10-030-226-1

Sequence: 1 gcaactctccgcgggtctctg.....tccagctctctctgaactac 2405

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estrov:*
6:	em_estrpl:*
7:	em_estro:*
8:	em_hic:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_hic:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	em_gss_hum:*
18:	em_gss_inv:*
19:	em_gss_pln:*
20:	em_gss_yrt:*
21:	em_gss_fun:*
22:	em_gss_mam:*
23:	em_gss_mus:*
24:	em_gss_pro:*
25:	em_gss_rdg:*
26:	em_gss_phg:*
27:	em_gss_vrl:*
28:	gb_gss1:*
29:	gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1335	55.5	1990	11	AK076014 Mus muscu
2	1308	54.4	2018	11	AK037042 Mus muscu
3	942.2	39.2	997	13	EX354148 BX354148
4	939.2	39.1	1201	13	EX460000 BX460000

5	917.6	38.2	1080	13	EX358662 BX358662
6	863	35.9	930	9	AL521329 AL521329
7	861.4	35.8	994	13	EX354149 BX354149
8	855	35.6	1200	13	EX353271 BX353271
9	839.6	34.9	908	13	BQ680322 BQ680322
10	837	34.8	870	9	AL529216 AL529216
11	833.6	34.7	934	9	AL521330 AL521330
12	786.8	32.7	903	13	BQ678639 BQ678639
13	774.8	32.2	973	9	AL529217 AL529217
14	767.2	31.9	1044	12	B1820310 B1820310
15	753.6	31.3	969	13	BX419446 BX419446
16	753.4	31.3	1067	12	EM559002 EM559002
17	734.6	30.5	1298	9	AV762219 AV762219
18	733.2	30.5	775	12	B1821631 B1821631
19	729.4	30.3	1178	13	BX353272 BX353272
20	728	30.3	827	13	B1915593 B1915593
21	721.8	30.0	911	13	B1950373 B1950373
22	717.4	29.8	837	12	BG720139 BG720139
23	705.2	29.3	975	13	BQ678464 BQ678464
24	682.4	28.4	711	12	BG719288 BG719288
25	678	28.2	679	12	BM727185 BM727185
26	673.6	28.0	895	13	BQ437413 BQ437413
27	664.2	27.6	928	13	BX412740 BX412740
28	658.4	27.4	992	13	B163419 B163419
29	658	27.4	790	12	B1819939 B1819939
30	658	27.4	811	10	BE798017 BE798017
31	645.4	26.8	1216	12	BM475414 BM475414
32	642.2	26.7	671	12	B1522488 B1522488
33	641.4	26.7	684	14	CD364420 CD364420
34	634.6	26.4	1105	12	BM451631 BM451631
35	629	26.2	669	12	BM678297 BM678297
36	625.8	26.0	1143	10	BE686280 BE686280
37	619.2	25.7	676	10	BE311800 BE311800
38	619	25.7	785	10	BE277176 BE277176
39	618.6	25.7	665	12	BM681655 BM681655
40	616.6	25.6	1040	10	BE543175 BE543175
41	616	25.6	663	13	B1625560 B1625560
42	615.4	25.6	905	14	CB194469 CB194469
43	610	25.4	796	10	BE274647 BE274647
44	607.4	25.3	828	12	B1910210 B1910210
45	598.8	24.9	905	13	B1519047 B1519047

ALIGNMENTS

RESULT 1
AK076014
LOCUS
DEFINITION Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:2610030019 product:solute carrier family 27
(fatty acid transporter), member 3, full insert sequence.

ACCESSION AK076014
VERSION AK076014.1
KEYWORDS GI:26344941
SOURCE HTCC, CAP trapper.
ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL
MEDLINE 98279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
JOURNAL
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Ishiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
20530913
11076661

REFERENCE
AUTHORS
4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

JOURNAL
AUTHORS
5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

JOURNAL
AUTHORS
6 (bases 1 to 1990)

JOURNAL
AUTHORS
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, K., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamoto, T., Yamahata, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES
source
1. 1990
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:2610030U19"
/db_xref="MGI:1908071"
/db_xref="taxon:10090"
/clone="2610030U19"
/tissue_type="whole body"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days embryo"
533..1873
/note="unnamed protein product; putative solute carrier family 27 (fatty acid transporter), member 3 (MCD|MGI:1347358, GB|AF072758, evidence: BLASTN, 99%, match=1387)";
/codon_start=1
/protein_id="BAC36120.1"
/db_xref="GI:26344942"

CDS

ORIGIN
polya_signal
polya_site
/note="putative"
1990
/note="putative"

Query Match 55.5%; Score 1335; DB 11; Length 1990;
Best Local Similarity 82.0%; Pred. No. 5,7e-217;
Matches 1583; Conservative 0; Mismatches 300; Indels 48; Gaps 2;

331 GGCCTTCGAGCTCGGCGCCCTGCGCGGCTGCGCGGAGAGTCCGAGGCGG 390
1 GGCCTTCGAGCGAGCGCGCTGCGCGGCGCGCGGAGAGTCCGAGGCGG 60
391 CTGAGCTGCGCGCTGCGCGCTGCGCGGAGAGTCCGAGGCGCGGAGAGT 450
61 GTGAGCTGCGCGCTGCGCGCTGCGCGGAGAGTCCGAGGCGCGGAGAGT 120
451 CATTCACGCGCTGCGCGCTGCGCGCTGCGCGGAGAGTCCGAGGCGCGGAG 510
121 CATTCACGCGCGCGCGAGGCTGCGCGGAGAGTCCGAGGCGCGGAGAGT 180
511 ACGCGCTTCCTGAGCTGCGCGTGAAGTCCGAGGAGAGTCCGAGGCGCGGAG 570
181 TCGCGCTTCCTGAGCTGCGCGTGAAGTCCGAGGAGAGTCCGAGGCGCGGAG 234
571 GGGAGAGCTGCGGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGAG 630
235 GGGAGAGCTGCGGAGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGAG 292
631 CGGCGGAGGAGTTCGCGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAG 690
293 -----GGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 312
691 GTACCTGAGAGAGTGTGCGCTGCTCCCGCTGCGCGGAGAGTTCGTGCTCTG 750
313 GGCACCGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 372
751 GTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 810
373 GTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 432
811 CCCCTGCTGAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 870
433 ACCCTGCTGAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 492
871 TCTGAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 930
493 CTGAGTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 552
931 TCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 990
553 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 612
991 AGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1050
613 AGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 672
1051 GTACATCTTCAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1110
673 GTACATCTTCAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 732
1111 GATCTGCAATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1170
733 GATCTGCAATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 792

QY	1111	CCCTGCGCCTCCCACTCTACACAAATGTCGGGTTCCTGCTGGGCATCGTGGGCTGATGAG	1230
Db	793	CTCCGACCTCCCACTGTAACAATGTCTGGCTCCTTCTGGGCATTTGGGCTGTTGGG	852
QY	1231	CATTGGGGCCACAGTGGTCTGAAATCCAAAGTTCGGCTGGTCAAGTCTGG3AAGATTG	1290
QY	1291	CCAGCAGCACAGGGTGAAGGTTCCAGTACATTGGGGAGCTGTGCCATACCTTGTCAA	1356
Db	913	CCAGAAACACAGGGGTGACAGTGTCCAGTACATTGGGGAGTGTGCCATACCTGTCAA	972
QY	1351	CCAGGCCCCGAGCAGGCAAGGCAAGACGTGGGCATTAAGGTCGGGCTGGCAGTGGGACGGGCT	1410
Db	973	CCAGCCCCCGAGCAGGCAAGGCAAGAGTGTGACCATTAAGGTGGCTTTGGCAGTGGGCACTGGGCTT	1032
QY	1411	GGGCCAAGTACCTGGGAGCGTTTGTGCGGCGCTTCGGGCCCCCTGCAGGTCTGAGAC	1470
Db	1033	GGGCCCAAGACACTGGGAGCGTTTCTCGGGGCAATTGGACCTCTGCAGATACCTGAGAC	1092
QY	1471	ATATGAGCTGACAGAGGGCAACGTGGGCCACCATCACTAACACAGGACAGCGGGGCGCTGT	1530
Db	1093	GTATGGCATGACAGGGCAACGTAGTACGTTCAATTACACAGGACGGGAGGGTGCAGT	1152
QY	1531	GGGGCGTCTCTCTGGCTTTTACAGACATATCTCCCTCTCTCTGATTCGCTATGATGT	1590
Db	1153	GGGGGAGCTTCTGTGGCTTTACAGACCATCTCCCTCTCTCTGATTCGCTATGATGT	1212
QY	1551	CACCAACAGAAAGCCAATTCCGGACCCCCCAAGGGGCACTGTATGGCACATCTCCAGGTGA	1650
Db	1213	CATGACAGAGGGGCGCTATTCCGAATGCCCAAGGGGCACTGACATGACACATCTCCAGGTGA	1272
QY	1651	GCAGAGGCTGTGGTGGGCGCCGGGTAAAGCAGAGAGTCCCATTCCTGGGCTATGGTGGGG	1710
Db	1273	GCGAGGCTTACTGGTGGGCCCAAGTAAAGCAAGTCTGAGTCTGGGCGATTTCTT	1332
QY	1711	GCCAGAGCTGGCCCAAGGGGAAAGTGTCTTAAGAGATGTCTTCGGGCTGGGATGTTTTCTT	1770
Db	1333	TCCGAGCTGGGCAAGGACCAAGCTGCTAAGAGATGTCTTCTGATCTGGGCGATTTCTT	1392
QY	1771	CAACACTGGGGGACCGTGGTCTGGGAGTACCAAGGTTTTTCCGCTTCCATGATCGAC	1830
Db	1393	CATATCTGGGACCTCTTGGTGTGATGAGCAGAGGCTTTCTTCACTTCCACGATCGAC	1452
QY	1831	TGAGACACTTCAGGTGGAAGGGGGAGATGTGGCCCAACCGAGGTGGCAGAGTCTT	1890
Db	1453	TGGAACACCTTCAGGTGGAAGGGGAGAGATCTGGCCCAACATGAAAGTGGTGAAGTCTT	1512
QY	1891	CGAGGCCCTAATTTTCTTCAGAGAGTAAAGTCTATGGAGTCACTGTGCCAGGGCATGA	1950
Db	1513	GGAGACCTTGACCTTCTTCAGAGAGTAAACATCTAGAGTCAACGGGCCAGGGGACGA	1572
QY	1951	AGGCAAGGCTGGAATGCGAGCCCTAGTCTCGTCCGCCCAAGGCTTTGACCTTATGCA	2010
Db	1573	AGGCAAGGCAAGGCAAGGCGGCTTGGCTGTGCGGCCCGGAGGCTCTGAACCTGTGTCA	1632
QY	2011	GCTCTACACCAACGTGTGTGAAGCTTGCACCTTATAGCCCGGCCCGGATTCCTAGGCT	2070
Db	1633	GCTCTACAGCCATGTTTCTGAGAACTTGCACCGATACCCGACCTCGGTTCTCAAGCT	1692
QY	2071	CCAGGAGCTTTGGGCAACACAGAGACTTCAACACAGAGAAAGTTGGGATGGCAATGA	2130
Db	1693	CCAGGAATCTTTGGGCACTACTGAGACCTTCAACACAGAGAAAGTTAGATGGCAATGA	1752
QY	2131	GGGCTTGACCCCAAGCACTGTGTGACCACTGTACGTTCTGGAACAAGCTGTAGTGC	2190
Db	1753	GGGCTTTGACCCCAAGTGTATGTGTGACCCACTATATTTCTGGAACAAGTATAGGGGC	1812
QY	2191	CTACCTGGCCCTCAACACTGGCCGGGTAAAGGGCCCTCTGGCAGAGAACTTGCATATCTG	2250
Db	1813	CTACCTGGCCCTCAACACTGGCCGGGTAAAGTGCCTCTGTCTGTGAGACTTGTGAATCTG	1872

QY	2251	AGACTTCAC	2261
Db	1873	AAACTTCAC	1883
RESULT 2	AK037042	2018 bp	mRNA
LOCUS	AK037042		linear
DEFINITION	Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930102C17 product:sequence carrier family 27 (fatty acid transporter), member 3, full insert sequence.		
ACCESSION	AK037042		
VERSION	AK037042.1		
KEYWORDS	HTC, CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1	Carninci, P. and Hayashizaki, Y.	
AUTHORS	High-efficiency full-length cDNA cloning		
TITLE	Meth. Enzymol. 303, 19-44 (1999)		
JOURNAL	9979253		
MEDLINE	10349666		
PIRBASE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.	
REFERENCE	3	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
AUTHORS	Genome Res. 10 (10), 1617-1630 (2000)		
TITLE	11042159		
JOURNAL	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishibe, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ilegami, T., Kashigagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matsuki, K., Yoneida, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format		
JOURNAL	sequencing pipeline with 384 multiplexillary sequencer		
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)		
PIRBASE	20530913		
AUTHORS	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
REFERENCE	5	Functional annotation of a full-length mouse cDNA collection	
TITLE	Nature 409, 685-690 (2001)		
JOURNAL	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
MEDLINE	Analysis of 60,770 full-length cDNAs		
PIRBASE	Nature 420, 563-573 (2002)		
AUTHORS	6 (bases 1 to 2018)		
TITLE	Adachi, J., Aizawa, K., Akinura, T., Aikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuta, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koyas, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.		
JOURNAL	Direct Submission		
MEDLINE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),		

COMMENT

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.

FEATURES

source

1. 2018
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM:DB:9930102C17"
 /db_xref="MGI:2401515"
 /db_xref="taxon:10090"
 /clone="9930102C17"
 /sex="female"
 /tissue_type="vaginal"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 <1.1901

misc_feature

/note="putative
 solute carrier family 27 (fatty acid transporter), member 3 (MGD) [MGI:1347358, GB|AF072758, evidence: BLASTN, 98%, match=1387]"
 2001..2006
 /note="putative"
 2018
 /note="putative"

ORIGIN

Query Match 54.4%; Score 1308; DB 11; Length 2018;
 Best Local Similarity 81.1%; Pred. No. 2.3e-212;
 Matches 1588; Conservative 0; Mismatches 300; Indels 70; Gaps 3;

QY 326 AAAAGGGCTCTTGAAGCTCGCGCCCTGCGCGGCTGCGCGGACCGGAAAGTCCCGAG 385
 Db 2 AAAAGGGCCCTTGAGCAAGCGCGCTGCGCGGACCGGACCGGAAAGTCTGAG 61
 QY 386 GGGGGCTGCAAGCTGCGCTGCGGCTCGCGGAACTGCGCGGACCGGACCGGACCG 445
 Db 62 AGCGGCTGCAAGTCTGCGCTGCGGCTCGCGGAACTGCGCGGACCGGACCGGACCG 121
 QY 446 TTTCCTATTCACGCGCTCGCGGCTTTAGCTACTAGAGCGGAGCGGAGCGGAAAGG 505
 Db 122 TTCTTATTCACGCGGCGGAGCGCTTTAGCTACCGGAGGCTGAGCGGAGGAGCAACGG 181
 QY 506 GCTGCAAGCGCTTCTCTACGTGCGCTAGGCTGAGGAACTGGGAACTGGGACCGGACCG 565
 Db 182 ATTGCTCGGCGCTTTTGTGCGGACACGCGGCTGAGACCGGAGG-----CGCGGAGGCTCG 235
 QY 566 GGGGAGGGAGGCGCTGGAAGAGGCGAGGCGGAGCGGAGCGGAGCGGAGATGACAGCGG 625
 Db 236 GAGAGGGAGGAGCTAGAGAGGCGACCGCTGCGGCTTCGCGCTGAGATCGGCTGCT 295
 QY 626 GGAAGCGGCGGAGTTTCCGAGAGGAGCGGTGCCCGCAAGTGAAGAGCGGCGGCC 685
 Db 296 AGA-----GGGAGCAACCGCGGCC 313
 QY 686 CCTCTGTACCTGGAACAACTGTGGGCTGTCTTCCCGGCGGCGGAGAGTTTGTGG 745
 Db 314 CCTCTGTACCTGGAACAACTGTGGGCTGTCTTCCAGCGGCGGAGATTCCTTGG 373
 QY 746 CTCTGTTGGGGCTGGGCAAGGCGGCTGCGGCACTGCTTTGTGCGCAACCGGCTGGCG 805
 Db 374 ATTGTGTTGGAGATGGCGCAAGAGTGGCTTGCGGAGGGCTTTGTGCGCAACCGCTTACGC 433
 QY 806 CGGGGCCCTGTGCTGCTCTCCGAGCTGCGGCGCGCGCGCTGTGTGCTGGCGCC- 864

Db 434 CGAGGACCCCTGTGTGCTGCTCCCGAGCTCGGAGGAGGCGGCTGTGCGGCA 493
 QY 865 -----AGGTTCTGAGTCCCTGGAGCGGAGCTGCGGCGCT 903
 Db 494 GCTGGTTTGTCTTACCTTGTGAGTCTCTGGAGTCTCTGGAGCGGAGCTTCCGCTT 553
 QY 904 GAGAGCATGGGGCTCACCTTGTGAGTCTCTGGAGTCTCTGGAGCGGAGCTTCCGCTT 563
 Db 554 GAGAGCATGGGGCTCACCTTGTGAGTCTCTGGAGTCTCTGGAGCGGAGCTTCCGCTT 613
 QY 964 CGATTGTGCTGCTGAGTCTCTGGAGTCTCTGGAGTCTCTGGAGCGGAGCTTCCGCTT 1023
 Db 614 CAATTGTCTATCGAAGCAGCAACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 673
 QY 1024 CCCCAGAGATACAGACACGCTGTATCTTACCTTGTGAGTGTGAGTGTGAGTGTGAG 1083
 Db 674 CCCCAGAGATACAGACACGCTGTATCTTACCTTGTGAGTGTGAGTGTGAGTGTGAG 733
 QY 1084 CAAAGCTGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1143
 Db 734 CAAAGCTGCTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 793
 QY 1144 TGGTGTCCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1203
 Db 794 TGGTGTCCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 853
 QY 1204 CTTGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1263
 Db 854 CTTGCTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 913
 QY 1264 CTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1323
 Db 914 CTGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 973
 QY 1324 TGGGAGCTGTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1383
 Db 974 TGGGAGCTGTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033
 QY 1384 GGTTCGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1443
 Db 1034 GGTTCGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1093
 QY 1444 CTTGCGGCTGTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1503
 Db 1094 ATTGGAACCTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1153
 QY 1504 CAATACACAGGACAGCGGAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 1563
 Db 1154 CAATACACAGGACAGCGGAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 1213
 QY 1564 CCGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1623
 Db 1214 CCGCTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1273
 QY 1624 GCACTGTATGACCAATCTTCAAGTGAAGCGGAGCTGTGGGCTGTGGGCTGTGGGCTGTGG 1683
 Db 1274 GCACTGTATGACCAATCTTCAAGTGAAGCGGAGCTGTGGGCTGTGGGCTGTGGGCTGTGG 1333
 QY 1684 GTCCCATTTCTGAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 1743
 Db 1334 GTCCCATTTCTGAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 1393
 QY 1744 TGTCTTCGAGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 1803
 Db 1394 TGTCTTCGAGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 1453
 QY 1804 AGGTTTCTTCCGCTTCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1863
 Db 1454 AGGTTTCTTCCGCTTCCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1513
 QY 1864 GGCACACACGAGTGTGAGGCTTCCGAGGCTTGAATTTTCTCAGAGGCTGAAGCT 1923

Db 1514 GGCCAACTGAAGTGGCTGAGGCTTGGAGACCCCTGGAATTCTTCAGAGGTGAACAT 1573
Qy 1924 CTATGAGACTGATGCCAGAGGATGAAGGAGGCTGGAAATGAGAGCCATGTTCTGCG 1983
Db 1574 CTATGAGTCAAGGTGCGAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1633
Qy 1984 TCCCCCAACGCTTGAAGCTTATGACGCTTACACCAAGTGTCTGAACCTTGCCACC 2043
Db 1634 GCGCCGAGGCTTGAAGCTTGTGACGCTTACAGCCATGTTCTGGAACCTTGCCACC 1693
Qy 2044 TTATGCG 2103
Db 1694 GTATGCG 1753
Qy 2104 ACAGCAGAAAGTTCGATGAGCAATGAGGCTTCAGACCCAGACCCCTGTCTGACCCACT 2163
Db 1754 ACAGCAGAAAGTTCGATGAGCAATGAGGCTTCAGACCCAGACCCCTGTCTGACCCACT 1813
Qy 2164 GTACGTTCTGAGCAGGCTGTAGTGCTTACCTGCGCTTACACCACTGCGCGGTACAGCGC 2223
Db 1814 CTATGTTCTGAGCAGGATATAGGGGCTTACCTGCGCGCTTACACCACTGCGCGGTACAGTG 1873
Qy 2224 CCTCTGCGAGGAACCTTCGATCTGAAGCTTCCAC 2261
Db 1874 CCTCTGCTGAGAGCTTCGATCTGAAGCTTCCAC 1911

RESULT 3
LOCUS BX354148 997 bp mRNA linear EST 05-MAY-2003
DEFINITION BX354148 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
ACCESSION BX354148
VERSION BX354148.1 GI:30373826
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 997)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 661.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DC015D1NP1&cluster=661.f. Contact :
Peng Liang Email: liliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DC015D1NP1.
Location/Qualifiers
source 1..997
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC015H21"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (4T)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 39.2%; Score 942.2; DB 13; Length 997;
Best Local Similarity 98.4%; Pred. No. 3.3e-150;
Matches 977; Conservative 6; Mismatches 6; Indels 4; Gaps 3;

Qy 1383 AGGTCGGCTGAGAGTGGGAGGCGGCTGCGCCCAAGATACCTGGAGACGCTTTGTGCGC 1442
Db 990 AAGTCCGGCTGAGAGTGGG--CAGCGCGGCCCAAGTACTGGAGCKTTTGTGCGC 993
Qy 1443 GCTTGGGCGCTTGAAGTGTCTGAGACATATGAACTGACAGAGGCAAGTGGCCCA 1502
Db 932 GCTTGGGCGCTTGAAGTGTCTGAGACATATGAACTGACAGAGGCAAGTGGCCCA 873
Qy 1503 TCAACTACAGAGACAGCGGCGGCTGTGGGCGTGTCTGCTTGAACATATCT 1562
Db 872 TCAACTACAGAGACAGCGGCGGCTGTGGGCGTGTCTGCTTGAACATATCT 813
Qy 1563 TCCCTTCTCTGATTCGCTATGATGACCAAGAGAGCCATTTGGGACCCCAAG 1622
Db 812 TCCCTTCTCTGATTCGCTATGATGACCAAGAGAGCCATTTGGGACCCCAAG 753
Qy 1623 GGCATGTATGAGCAATCTCCAGGTGAGCGAGGCTGTGTGGCCCGGTAAGCCAGC 1682
Db 752 GGCATGTATGAGCAATCTCCAGGTGAGCGAGGCTGTGTGGCCCGGTAAGCCAGC 693
Qy 1683 AGTCCCATTCCTGGGCTATGCTGCGGCGCAGAGCTGGCCAGGGAAATTCTAAAG 1742
Db 692 AGTCCCATTCCTGGGCTATGCTGCGC--GGCAGAGCTGGCCAGGGAAATTCTAAAG 634
Qy 1743 ATGCTTCCGGCTGGGATGTTTCTTCAACACTGGGGACTGCTGTCTGATGAC 1802
Db 633 ATGCTTCCGGCTGGGATGTTTCTTCAACACTGGGGACTGCTGTCTGATGAC 574
Qy 1803 AAGGT-TTTCGCGCTTCCATGATGCTGAGAGACCTGAGGTGAAGGGGGGAAT 1861
Db 573 AAGGTATTTCTCGCTTCCATGATGCTGAGAGACCTGAGGTGAAGGGGGGAAT 514
Qy 1862 GTGGCCCAACCGAGGTGGGAGAGTCTTGAGGCGCTTGAATTTCTTGAAGAGTGAAC 1921
Db 513 GTGGCCCAACCGAGGTGGGAGAGTCTTGAGGCGCTTGAATTTCTTGAAGAGTGAAC 454
Qy 1922 GTCTATGAGTCACTGTGCGAGGCGATGAAGGAGGCTTGAATGGCAACCTTCTG 1981
Db 453 GTCTATGAGTCACTGTGCGAGGCGATGAAGGAGGCTTGAATGGCAACCTTCTG 394
Qy 1982 CGTCCCGCCAGCGCTTGAAGCTTATGAGAGCTTACACCAAGTGTCTGAGAACTTGCA 2041
Db 393 CGTCCCGCCAGCGCTTGAAGCTTATGAGAGCTTACACCAAGTGTCTGAGAACTTGCA 334
Qy 2042 CTTTATGCG 2101
Db 333 CTTTATGCG 274
Qy 2102 AACACAGAGAAATGCGATGGCAATGAGGCGTTGAGCCCGAGCAACCTGTGTACCA 2161
Db 273 AACACAGAGAAATGCGATGGCAATGAGGCGTTGAGCCCGAGCAACCTGTGTACCA 214
Qy 2162 CTGTAGCTTGTGACAGGAGCTGTAGGTGCTTACCTGCGCTTCAACCTGCGCGTACAGC 2221
Db 213 CTGTAGCTTGTGACAGGAGCTGTAGGTGCTTACCTGCGCTTCAACCTGCGCGTACAGC 154
Qy 2222 GCGCTCCTGAGAGAACTTGAATCTGAGAACTTCCACACTGAGGAGCACTGAGAGAG 2281
Db 153 GCGCTCCTGAGAGAACTTGAATCTGAGAACTTCCACACTGAGGAGCACTGAGAGAG 94
Qy 2282 GAATCTGTGGGTGGGCGCGGTGAGGTGTATGAGGCTGTCAAGGATCTTTCTATAC 2341
Db 93 GAATCTGTGGGTGGGCGCGGTGAGGTGTATGAGGCTGTCAAGGATCTTTCTATAC 34
Qy 2342 CAGAATGTGGTCACTATTTTGAATTAATGTG 2374
Db 33 CAGAATGTGGTCACTATTTTGAATTAATGTG 1

RESULT 4
LOCUS BX460000 1201 bp mRNA linear EST 22-MAY-2003
DEFINITION BX460000 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone

CS0DF009YF05 5-PRIME, mRNA sequence.
 ACCESSION BX460000
 VERSION BX460000.1 GI:31023191
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1201)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 661.f For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DF009C03P1&cluster=661.f. Contact :
 Feng Liang Email : fliang@life-tech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DF009C03P1.

FEATURES
 source
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF009YF05"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_1lb="Homo sapiens FETAL BRAIN"
 /note="Organ: Brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 39.1%; Score 939.2; DB 13; Length 1201;
 Best Local Similarity 96.4%; Pred. No. 1.2e-149;
 Matches 972; Conservative 8; Mismatches 26; Indels 2; Gaps 2;
 QY 35 GGAGTGGTGGGGCCCGGATGGGATGGGCTGGTGGCGAGCGCGGCTCCCTGGAAG 94
 DB 43 GTHCCGATCCGGAATTCGCCGATATGAGCGCTGTGCGACGACCGGGCTCCCTGGAAG 102
 QY 95 GAGAGCTCTAGCTAGAAACGAGCGCGCTGATTTTCGGAAGGAGATCAGGATGTTT 154
 DB 103 GAGAACTCTAGCTAGAAACGAGCGCGCTGATTTTCGGAAGGAGATCAGGATGTTT 162
 QY 155 GCGAGGCGCTGGAACCGAGCGGTCGCGATGAGGAGGAGGCGGCTCGATGGCTGCTCTG 214
 DB 163 GCGAGGCGCTGGAACCGAGCGGTCGCGATGAGGAGGAGGCGGCTCGATGGCTGCTCTG 222
 QY 215 CTGCTGCCCCCTGCTGTGTGCTACCGCTGCTGCTGTAAGCTACACCTTGGCCGAG 274
 DB 223 CTGCTGCCCCCTGCTGTGTGCTACCGCTGCTGCTGTAAGCTACACCTTGGCCGAG 282
 QY 275 TTGCGCTGGCTTCCCGCGAATTGCGCTTGGCGGCGAGCTGTGCTGCGAAAGGGCT 334
 DB 283 TTGCGCTGGCTTCCCGCGAATTGCGCTTGGCGGCGAGCTGTGCTGCGAAAGGGCT 342
 QY 335 CTTCGAGCTCGCGCCCTGGCGCGGCTGCGCGCGAAGCTCCGAGGGGGGCTGC 394
 DB 343 CTTCGAGCTCGCGCCCTGGCGCGGCTGCGCGCGAAGCTCCGAGGGGGGCTGC 402
 QY 395 AGCTTGGCTGGCTGCGGCTGCGGCTGCGCGCGAAGCTCCGAGGGGGGCTGC 454
 DB 403 AGCTTGGCTGGCTGCGGCTGCGGCTGCGCGCGAAGCTCCGAGGGGGGCTGC 462
 QY 455 CAGGCTCGGCGGCTTGTAGCTACTGAGAGGCGGAGCGGAGTAACAGGGCTGACGC 514

DB 463 CAGGCTCGGCGGCTTGTAGCTACTGAGAGGCGGA-SGGAAGTAACAGGGCTGACCC 521
 QY 515 GCCTTCCTACGTCGCTGAGCTGGAGCTGGAGACCGGAGCGGAGCGGAGGGG 574
 DB 522 GCCTTCCTACGTCGCTGAGCTGGAGCTGGAGACCGGAGCGGAGCGGAGGGG 581
 QY 575 AGGCTTGAAGAGGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 634
 DB 582 AGGCTTGAAGAGGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 641
 QY 635 GCGAGTTTCCCGAGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 694
 DB 642 GCGAGTTTCCCGAGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 701
 QY 695 CTTGAGCAACTGTGCGCTGCTCTCTCCCTGCGGAGGAGTTTGTGTGCTTGTTC 754
 DB 702 CTTGAGCAACTGTGCGCTGCTCTCTCCCTGCGGAGGAGTTTGTGTGCTTGTTC 761
 QY 755 GGGCTGGCAAGGCGGCGGCTGCGCACTGCTTTGTGCCACCGCCCTGCGGCGGCGCC 814
 DB 762 GGGCTGGCAAGGCGGCGGCTGCGCACTGCTTTGTGCCACCGCCCTGCGGCGGCGCC 821
 QY 815 CTGCTCACTGCTCTGCGAGCTGCGGCGGCGGCGGCTGCTGCGGCGGAGATTTCG 874
 DB 822 CTGCTCACTGCTCTGCGAGCTGCGGCGGCGGCGGCTGCTGCGGCGGAGATTTCG 881
 QY 875 GATTCCTGAGAGCGGAGCTGCGGCGGCGGCGGCTGAGAGGAGGAGGAGGAGGAGG 934
 DB 882 GATTCCTGAGAGCGGAGCTGCGGCGGCGGCGGCTGAGAGGAGGAGGAGGAGGAGG 941
 QY 935 GGGCCAGAACCCACCTGCTGGAATTAAGATTGCTGCTGAGTGTCC-GCTGAGT 993
 DB 942 GGCAGAGAACCCACCTGCTGGAATTAAGATTGCTGCTGAGTGTCCGCTGAGT 1001
 QY 994 GATGGGCGAGTCCGAGTACTCTCTTCCCTCCCGGAGGAGGAGGAGGAGGAGG 1041
 DB 1002 GATGGGCGAGTCCGAGTACTCTCTTCCCTCCCGGAGGAGGAGGAGGAGGAGG 1049

RESULT 5

EX358662 1080 bp mRNA linear EST 05-MAY-2003
 LOCUS EX358662 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1042YK09 5-PRIME, mRNA sequence.
 ACCESSION EX358662
 VERSION EX358662.1 GI:30380277
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1080)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 661.f For more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D1042AF05Q1&cluster=661.f. Contact :
 Feng Liang Email : fliang@life-tech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0D1042AF05Q1.
 FEATURES
 source
 1..1080
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

ORIGIN

/clone="CS0D1042YK09"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 38.2%; Score 917.6; DB 13; Length 1080;
Best Local Similarity 97.8%; Pred. No. 5.3e-146;
Matches 939; Conservative 10; Mismatches 9; Indels 2; Gaps 2;

```

QY 11 CCGGTTTCTGCTCTCCGCGCGCTGAGTGTGGGGGCTGGTGGAGTGGGCGCTGGC 70
DB 64 CCGGTTTCTGCTCTCCGCGCGCTGAGTGTGGGGGCTGGTGGAGTGGGCGCTGGC 123
QY 71 CAGCCGACGCGCGCTCTCCGCGCGCTGAGTGTGGAGTGTGGAGTGTGGAGTGT 130
DB 124 CAGCCGACGCGCGCTCTCCGCGCGCTGAGTGTGGAGTGTGGAGTGTGGAGTGT 182
QY 131 CCGAAGGGAGATCAAGGATGTTTGGAGGCGCTGGAACCGAGCGTGGCAGTGAAGAA 190
DB 183 CCGAAGGGAGATCAAGGATGTTTGGAGGCGCTGGAACCGAGCGTGGCAGTGAAGAA 242
QY 191 GCGGCGCTCCATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 250
DB 243 GCGGCGCTCCATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
QY 251 CTGAAGCTACACCTCTGCGCGCGAGTGTGGGCTGCTGCGGAGTGTGGGCTGCTG 310
DB 303 CTGAAGCTACACCTCTGCGCGCGAGTGTGGGCTGCTGCGGAGTGTGGGCTGCTG 362
QY 311 CGAGCTCTGTGTGCAAAAGGAGCTCTTCAAGCTGCGCGCTGCGCGCGCTGCGCG 370
DB 363 CGAGCTCTGTGTGCAAAAGGAGCTCTTCAAGCTGCGCGCTGCGCGCGCTGCGCG 422
QY 371 CCGGAAGGTCCCGAGGCGGCGCTGCAAGCTGCGCGCTGCGCGCGCTGCGCGAG 430
DB 423 CCGGAAGGTCCCGAGGCGGCGCTGCAAGCTGCGCGCTGCGCGCGCTGCGCGAG 482
QY 431 CCGCGCGCGCAACCTTCTCAATTCAGAGGCTGCGGCGCTTCAAGTACTGAGAGCG 490
DB 483 CCGCGCGCGCAACCTTCTCAATTCAGAGGCTGCGGCGCTTCAAGTACTGAGAGCG 541
QY 491 CCGGAGACTTACAGGAGCTGCAAGCGCTTCTCAAGTGTGCTGAGCTGAGGAGT 550
DB 542 CCGGAGACTTACAGGAGCTGCAAGCGCTTCTCAAGTGTGCTGAGCTGAGGAGT 601
QY 551 GACGCGCGCGCAAGGCGGAGGAGGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAG 610
DB 602 GACGCGCGCGCAAGGCGGAGGAGGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAG 661
QY 611 GAGAGTGCAGGCGCGGAGGAGGAGGAGTGTGCGGAGGAGGAGGAGGAGGAGT 670
DB 662 GAGAGTGCAGGCGCGGAGGAGGAGGAGTGTGCGGAGGAGGAGGAGGAGGAGT 721
QY 671 GAGAGGAGGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
DB 722 GAGAGGAGGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781
QY 731 CCAAGATTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 790
DB 782 CCAAGATTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841
QY 791 CCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
DB 842 CCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
QY 851 CTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 910
DB 902 TGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961

```

QY 911 ATGGGGCTCCACCTGTCGCGCTGAGGCGCCAGAGACCAACCTGCTGGAATTACGATTG 970
DB 962 ATGGGGCTCCACCTGTCGCGCTGAGGCGCCAGAGACCAACCTGCTGGAATTACGATTG 1021

RESULT 6
AL521329/c
DEFINITION
AL521329 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
LOCUS
AL521329
ACCESSION
AL521329
VERSION
AL521329.2 GI:31039632
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
On Feb 13, 2001 this sequence version replaced gi:12784822.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 661.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DB001BC09NP1cluster=661.f. Contact :
Peng Liang Email : filiang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DB001BC09NP1.

FEATURES

Source

1..930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DB001YE18"
/tissue_type="NEUROBLASTOMA COT 10-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 35.9%; Score 863; DB 9; Length 930;
Best Local Similarity 97.0%; Pred. No. 9.9e-137;
Matches 902; Conservative 16; Mismatches 8; Indels 4; Gaps 4;

```

QY 1453 CCTGCAAGTGTGAGACATATGACTGACAGAGGAGCAAGTGGCCACATCACTAC 1512
DB 930 CCTGCAAGTGTGAGACATATGACTGACAGAGGAGCAAGTGGCCACATCACTAC 871
QY 1513 AGG-AACAGCGGGCGCTGTGGGCGCTGCTTCTGCTTAAAGCAATCTTCCCTCT 1571
DB 870 AGGAAACAGCGGGCGCTGTGGGCGCTGCTTCTGCTTAAAGCAATCTTCCCTCT 811
QY 1572 CCTGATTCGCTATGATGTCACACACAGAGAGCAATGGAGACCCCGAGGGGCACTGA 1631
DB 810 CCTGATTCGCTATGATGTCACACACAGAGAGCAATGGAGACCCCGAGGGGCACTGA 751
QY 1632 TGGCAACATCTCCAGGTGAGCCAGAGGCTGCTGTGGCGCCCGGTAAAGCAAGTCCCAT 1691
DB 750 TGGCAACATCTCCAGGTGAGCCAGAGGCTGCTGTGGCGCCCGGTAAAGCAAGTCCCAT 691
QY 1692 TCCCTGGGCTATGCTGGCGGGGCGAGAGCTGGCCAGGGGAGAGTGTCTAAAGATGCTTCC 1751
DB 690 TCCCTGGGCTATGCTGGCGGGGCGAGAGCTGGCCAGGGGAGAGTGTCTAAAGATGCTTCC 632
QY 1752 GGCCTGGGAGATGTTTCTTCAACACTGGGAGCTGCTGCTGCGATGACCAAGGTTTTC 1811

```

Db	631	GGCCTGGGGAGTGTTTTCTTTCACCAACTGGGGAGACCTGCTGCTCGCATGACCAAGGTTTTC	572
QY	1812	TCCGCTTCCATGATGTGTAACCTGGAGACACCTTCAGGTGGAAGGGGAGAAATGTGGCCACAA	1871
Db	571	TCCGCTTCCATGATGTGTAACCTGGAGACACCTTCAGGTGGAAGGGGAGAAATGTGGCCACAA	512
QY	1872	CCGAGGTGGCAGAGGCTCTTCAGAGCCCTAGATTTTCTTCAGAGAGTGAACGTCTATGAG	1931
Db	511	CCGAGGTGGCAGAGGCTCTTCAGAGCCCTAGATTTTCTTCAGAGAGTGAACGTCTATGAG	452
QY	1932	TCACGTGCCCAGGGCATGAAGGCCAGGGCTGGAATGCAAGCCCTAGTTCTTGCGTCCCCC	1991
Db	451	TCACGTGCCCAGGGCATGAAGGCCAGGGCTGGAATGCAAGCCCTAGTTCTTGCGTCCCCC	392
QY	1992	ACGCTTGGACCTTATGACGCTCTCTACCCACGCTCTGAGAACTGCGCACCTTATGCCC	2051
Db	351	ACGCTTGGACCTTATGACGCTCTCTACCCACGCTCTGAGAACTGCGCACCTTATGCCC	332
QY	2052	GGCCCCGATTCCTCAGGCTCCAGAGGTCTTTGGCCACACAGAGACCTTCAACAGCAGGA	2111
Db	331	GGCCCCGATTCCTCAGGCTCCAGAGGTCTTTGGCCACACAGAGACCTTCAACAGCAGGA	272
QY	2112	AAGTTCGATGGCAATGAGAGGCTTCGACCCAGACCCCTGCTCAACCACTGTACGTTTC	2171
Db	271	AAGTTCGATGGCAATGAGAGGCTTCGACCCAGACCCCTGCTGTACCCACTGTACGTTTC	212
QY	2172	TGGACAGAGCTGTAGGTGCTTACCTGCCCCCTTCAACACTGCCCGGTACAGCCGCTTCCTGG	2231
Db	211	TGGACAGAGCTGTAGGTGCTTACCTGCCCCCTTCAACACTGCCCGGTACAGCCGCTTCCTGG	152
QY	2232	CAGGAAACCTTGAACTGTGAGAACTTCACACCTTAGGCACTGTGAGAGAGAACTCTGTG	2291
Db	151	CAGGAAACCTTGAACTGTGAGAACTTCACACCTTAGGCACTGTGAGAGAGAACTCTGTG	92
QY	2292	GGGTGGGGGCCCTTGCAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACCAAGAACTGGC	2351
Db	91	GGGTGGGGGCCCTTGCAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACCAAGAACTGGC	33
QY	2352	GTCACATATTTTGTATTAATGATGGGTGGAG	2381
Db	32	GTASATGCTTTT-TAATTAATGTGGCGGAG	4

RESULT	7
LOCUS	BX354149
DEFINITION	994 bp mRNA linear EST 05-MAY-2003
ACCESSION	BX354149 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
VERSION	CDNA Clone CS0DC015YH21 5-PRIME, mRNA sequence.
KEYWORDS	BX354149 BX354149 GI:30375843
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1. (bases 1 to 994)
COMMENT	Li W.B., Gruber C., Jesse J. and Polayes D. Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 661.f For more information about this cluster, see http://www.genoscope.cns.fr/ http://cgl.bim.cluster.cgi?seq=CS0DC015CD11QPI&cluster=661.f . Contact : Peng Liang Email : liang@life tech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope ID : CS0DC015CD11QPI.
FEATURES	Location/Qualifiers

```

source
1. .994
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CSDDC015H21"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI -0.190 (GT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

Query Match	Similarity	95.8%;	Score	861.4;	DB	13;	Length	994;
Best Local	Similarity	96.9%;	Pred.	1.9e-136;				
Matches	918;	Conservative	11;	Mismatches	12;	Indels	6;	Gaps
QY	24	TCGCGCCGCTGTGGAGTGCTGGGAGGCGCTGTGGTGGGATGGGCGTGTGCCATGCAACGCGCG	83					
DB	52	TCGGCAATATCCCGGAGATGTGGGGGGCGCTGGGTGGGATGGGGTGTGGCCAGCGCAACGGCGG	111					
QY	84	CTCCCTCGGAAGGAGAGTCTCAGCTTAAACGAGCGCGCCCTTAGTGTGTTTGGAAAGGAGAGAT	143					
DB	112	CTCCCTCGGAAGGAGAGTCTCAGCTTAAACGAGCGCGCCCTTAGTGTGTTTGGAAAGGAGAGAT	171					
QY	144	CAGGAGATGTTTGGAGCGGCTGGAACCAACGCGTGCAGATAGAGAAACCGGCTTCAATGG	203					
DB	172	CAGGAGATGTTTGGAGCGGCTGGAACCAACGCGTGCAGATAGAGAAACCGGCTTCAATGG	231					
QY	204	CTGCGCTCTCTGCTGCTGCGCCCTGCTGCTGCTGCTAAGCGCTGCTGCTGCTGCTGCTGCTAACC	265					
DB	232	CTGCGCTCTCTGCTGCTGCGCCCTGCTGCTGCTGCTAAGCGCTGCTGCTGCTGCTGCTGCTAACC	291					
QY	264	TCTGGCGCGAGTTGGCCCTGGCTTCCGGCGGACCTTGGCGGTGGAGCTCTGTGCT	322					
DB	292	TCTGGCGCGAGTTGGCCCTGGCTTCCGGCGGACCTTGGCGGTGGAGCTCTGTGCT	351					
QY	324	GCAAAAGGCTCTTTCAGCTCGCGCCCTTGGCGCGGCTGCGCGGACCCCGGAAGGTCOCG	383					
DB	352	GCAAAAGGCTCTTTCAGCTCGCGCCCTTGGCGCGGCTGCGCGGACCCCGGAAGGTCOCG	411					
QY	384	AGGGGGGCTGACGCTGCGGCTGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGA	443					
DB	412	AGGGGGGCTGACGCTGCGGCTGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGA	471					
QY	444	CCTTTCTCATTCACGGCTGCGCGCGCTTATGCTACTCAGAGCGGAGCTGCGAGATTAACA	503					
DB	472	CTTTTCTCATTCACGGCTGCGCGCGCTTATGCTACTCAGAGCGGAG- GCGAGATTAACA	530					
QY	504	GCGCTGACCGCGCTTCTTAACGTGCGCTGAGCTGAGGACCTGGGAGACCGGACGCGCGACA	565					
DB	531	GCGCTGACCGCGCTTCTTAACGTGCGCTGAGCTGAGGACCTGGGAGACCGGACGCGCGAMA	593					
QY	564	GCGCGAGGAGGACGCTGAGGAGGAGGAGCGGCGGACCGCGGAGCTGAGGATGCAAGCGG	623					
DB	591	-CGCGAGGAGGACGCTGAGGAGGAGGAGCGGCGGACCGCGGAGCTGAGGATGCAAGCGG	649					
QY	624	CCGGAACCGCGCGGAGCTTTCGCGAGGAGGAGGAGCTGCGGACGAGGTGGAAGACCGCGG	688					
DB	650	CCGGAACCGCGCGGAGCTTTCGCGAGGAGGAGGAGCTGCGGACGAGGTGGAAGACCGCGG	705					
QY	684	CCCTCTGTACCTGAGGACACTGTATGGCGCTCTCTCCCGCTGGGCCAGAGTTTCTGT	743					
DB	710	CCCTCTGTACCTGAGGACACTGTATGGCGCTCTCTCCCGCTGGGCCAGAGTTTCTGT	765					
QY	744	GCGCTGTGTTGGGCTTGGCCAGAGCGGCGCTTGGCACTGCTTTTGTGCCACCGCTTGC	803					
DB	770	GCGCTGTGTTGGGCTTGGCCAGAGCGGCGCTTGGCACTGCTTTTGTGCCACCGCTTGC	822					
QY	804	GCGCGGCGCGCGCTGCTGCACTGCTCTCGGAGCTGCGGCGCGCGCGCTGCTGCTGCGG	867					
DB	830	GCGCGGCGCGCGCTGCTGCACTGCTCTCGGAGCTGCGGCGCGCGCGCTGCTGCTGCGG	883					

QY 863 CCAGATTTCAGATCCCTGGAGCGGAGCCTGCCCCCTGAGAGCCAGTGGGCTCCAC 922
 Db 890 CCARATTTTCAGATCCCTGGAGCGGAGCCTGCCCCCTGAGAGCCAGTGGGCTCCAC 948
 QY 923 CTGTGGGCTGAGCGGCCAGGAACCACTCTGTGGAATTAAGCATTT 969
 Db 949 CTGTGGGCTGAGCGGCCAGGAACCACTCTGTGGAATTAAGCATTT 993

RESULT 8
 LOCUS BX353271/1
 DEFINITION BX353271 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DC003YF17 3-PRIME, mRNA sequence.
 ACCESSION BX353271
 VERSION BX353271.1 GI:30371745
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1200)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 661.f for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DC003CC09NP1&cluster=661.f. Contact :
 Feng liang Email : fliang@life.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0DC003CC09NP1.

FEATURES
 source
 location/Qualifiers
 1..1200
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC003YF17"
 /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="First strand cDNA was primed with a NcoI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 35.6% Score 855; DB 13; Length 1200;
 Best Local Similarity 97.1%; Pred. No. 2.5e-15;
 Matches 891; Conservative 9; Mismatches 15; Indels 3; Gaps 3;

Db 681 CTGGCGGCGCAGAGCTGGCCAGGGAGATGTCTTAAGATGTCTTCGGGCTGGGAGATG 622
 QY 1764 TTTTCTTCAACATGGGAGACCTGTGTGGATGACCAAGGTTTCTCCGCTTCATG 1823
 Db 621 TTTTCTTCAACATGGGAGACCTGTGTGGATGACCAAGGTTTCTCCGCTTCATG 562
 QY 1824 ATGTACTGAGACACCTTCAGGTGAGAGGGAGAAATGTGGCCACACGAGTGGCAG 1883
 Db 561 ATGTACTGAGACACCTTCAGGTGAGAGGGAGAAATGTGGCCACACGAGTGGCAG 502
 QY 1884 AGTCTTTCAGAGGCGCCCTAGATTTTCTTTCAGAGGTGAACTCTATGAGTCACTGTGCGCAG 1943
 Db 501 AGTCTTTCAGAGGCGCCCTAGATTTTCTTTCAGAGGTGAACTCTATGAGTCACTGTGCGCAG 442
 QY 1944 GGCATGAGGCGAGGCTGGATATGAGCAGCCCTAGTTGTGCTCCGCCCAAGCTTTGAGAC 2003
 Db 441 GGCATGAGGCGAGGCTGGATATGAGCAGCCCTAGTTGTGCTCCGCCCAAGCTTTGAGAC 382
 QY 2004 TTATGACGCTCTACACCCACAGTGTCTGAGAACTTGTGCACTTATGCGGCGCCGATTC 2063
 Db 381 TTATGACGCTCTACACCCACAGTGTCTGAGAACTTGTGCACTTATGCGGCGCCGATTC 322
 QY 2064 TCAGGCTCCAGAGGTCTTTGGCCACACAGAGACCTTAAACAGAGAAAGTTGGATG 2123
 Db 321 TCAGGCTCCAGAGGTCTTTGGCCACACAGAGACCTTAAACAGAGAAAGTTGGATG 262
 QY 2124 CAATAGAGGCTTGCAGCCCGACCTGTCTGACCACTGATGCTTGGACAGGCTG 2183
 Db 261 CAATAGAGGCTTGCAGCCCGACCTGTCTGACCACTGATGCTTGGACAGGCTG 202
 QY 2184 TAGTGTCTTACCTGCGCCCTTCAACACTGCGGAGACGCGCCCTCGGAGAGAACTCTG 2243
 Db 201 TAGTGTCTTACCTGCGCCCTTCAACACTGCGGAGACGCGCCCTCGGAGAGAACTCTG 142
 QY 2244 GAATCTGAGAACTTCCACACTGAGGACCTGAGAGAACTCTGTGGGCTGGGCGCG 2303
 Db 141 GAATCTGAGAACTTCCACACTGAGGACCTGAGAGAACTCTGTGGGCTGGGCGCG 82
 QY 2304 TTGCAGGTGATCTGGGCTGTGAGGATCTTTTCTATACAGAACTGCGGTACTATTTTG 2363
 Db 81 TTGCAGGTGATCTGGGCTGTGAGGATCTTTTCTATACAGAACTGCGGTACTATTTTG 22
 QY 2364 TAATAATGTGGCTGGAG 2381
 Db 21 TAATAATGTGGAGAG 4

RESULT 9
 LOCUS B0680322
 DEFINITION AGENCOURT 8032839 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6091179
 5', mRNA sequence.
 ACCESSION B0680322
 VERSION B0680322.1 GI:21793001
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 908)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@b6-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2333 row: m column: 04

Db 690 GTGGCCCGGTAAGCAGCACTCCCATTCCTGGGCTATGCTGGC-GGCCAGAGCTGACC 632
 QY 1724 CAGGGGAAGTGTCTAAAGAGATGTCTCCGGCTGGGGATGTTTCTTCAACACTGGGAC 1783
 Db 631 CAGGGGAAGTGTCTAAAGAGATGTCTCCGGCTGGGGATGTTTCTTCAACACTGGGAC 572
 QY 1784 CTGGCTGCTCCGATGACCAAGTCTTCTCCGCTTCCATGATGCTACTGAGACACTTC 1843
 Db 571 CTGGCTGCTCCGATGACCAAGTCTTCTCCGCTTCCATGATGCTACTGAGACACTTC 512
 QY 1844 AGTGAAGGGGAGATGTG3CCACAACCGAGGTGCGAGAGTCTTCAGAGCCCTAGAT 1903
 Db 511 AGTGAAGGGGAGATGTG3CCACAACCGAGGTGCGAGAGTCTTCAGAGCCCTAGAT 452
 QY 1904 TTTCTTCAGAGAGTGAAGTCTATGAGTCACTGTGCCAGGGCATGAAAGCAGGGCTGGA 1963
 Db 451 TTTCTTCAGAGAGTGAAGTCTATGAGTCACTGTGCCAGGGCATGAAAGCAGGGCTGGA 392
 QY 1964 ATGGAGCCCTAGTCTGCTGCCGCCGCCCGCTTGAAGCTTATGAGCTTACACCCAC 2023
 Db 391 ATGGAGCCCTAGTCTGCTGCCGCCGCCCGCTTGAAGCTTATGAGCTTACACCCAC 332
 QY 2024 GTGTCTGAAGACTTCCACTTATATGCCGCCCGGCTTCTCAGGCTTCAGAGCTTTTG 2083
 Db 331 GTGTCTGAAGACTTCCACTTATATGCCGCCCGGCTTCTCAGGCTTCAGAGCTTTTG 272
 QY 2084 GCCACCAAGAGACCTTCAACAGAGAAAGTGGATGAGCAATGAGGGCTTCGACCC 2143
 Db 271 GCCACCAAGAGACCTTCAACAGAGAAAGTGGATGAGCAATGAGGGCTTCGACCC 212
 QY 2144 AGCACCCTGTCTGACCCACTGTAGCTTCTGAGCAGAGCTGAGTGTGCTACTGCCCC 2203
 Db 211 AGCACCCTGTCTGACCCACTGTAGCTTCTGAGCAGAGCTGAGTGTGCTACTGCCCC 152
 QY 2204 ACAACTGCCCGGTACAGGCGCTCTCTGCGAGAGAACTTGAATCTGAACTTCCAC 2263
 Db 151 ACAACTGCCCGGTACAGGCGCTCTCTGCGAGAGAACTTGAATCTGAACTTCCAC 92
 QY 2264 CTGAGGCACTGAGAGAGAACTCTGTGGGCTGGGGCGCTGAGGATGATGAGGCTGT 2323
 Db 91 CTGAGGCACTGAGAGAGAACTCTGTGGGCTGGGGCGCTGAGGATGATGAGGCTGT 32
 QY 2324 CAGGATTTTCTATACCAAGACTCGGCTC 2354
 Db 31 CAGGATTTTCTATACCAAGATMGCTM 1
 RESULT 11
 AL521330 934 bp mRNA linear EST 22-MAY-2003
 LOCUS AL521330 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
 DEFINITION CDNA clone CS0DB001YE18 5-PRIME, mRNA sequence.
 ACCESSION AL521330
 VERSION AL521330.2 GI:31039633
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 934)
 L1.W.B., Gruber, C., Jesses, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12784823.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 661.f. For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DB001BC09Q0Pikcluster=661.f. Contact :

Feng Liang Email: fliang@life.techn.com URL:
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0DB001BC09Q0P1.
 Location/Qualifiers
 1. 934
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DB001YE18"
 /issue="NEUROBLASTOMA COT 10-NORMALIZED"
 /clone="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
 /note="Left strand cDNA was primed with a NotI-clisodIT primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 34.7%; Score 833.6; DB 9; Length 934;
 Best Local Similarity 98.6%; Pred. No. 1e-11; Indels 2; Gaps 2;
 Matches 858; Conservative 3; Mismatches 7;

QY 11 CGGGTTTCTGCTCTCCGCCCGTGTGAGTGTGGGGGCTG3GTGGAATGG3CGTGTGC 70
 Db 67 CGGGTTTCTGCTCTCCGCCCGTGTGAGTGTGGGGGCTG3GTGGAATGG3CGTGTGC 126
 QY 71 CAGGCAAGCGGGCTCCCTGGAAGAGAGTCTCAGCTGGAACGAGCGGCCCTAGGTTT 130
 Db 127 CAGGCAAGCGGGCTCCCTGGAAGAGAGTCTCAGCTGGAACGAGCGGCCCTAGGTTT 186
 QY 131 CCGAAGGAGAGATCAAGGATGTTTGGAGCGGCTGGAACAGACGGTGCATAGAGAA 190
 Db 187 CCGAAGGAGAGATCAAGGATGTTTGGAGCGGCTGGAACAGACGGTGCATAGAGAA 246
 QY 191 GCGGGCTCCATGGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
 Db 247 GCGGGCTCCATGGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306
 QY 251 CTGAAGCTACACCTCTGGCGCGAGTTGCGCTGAGCTCCGCGGAGACTTGGCGGTG 310
 Db 307 CTGAAGCTACACCTCTGGCGCGAGTTGCGCTGAGCTCCGCGGAGACTTGGCGGTG 366
 QY 311 CCGAGCTGTGCTGCAAAAGGCTCTTGAAGTGTGGGCTGCGGCTGCGGCTGCGGCGAG 370
 Db 367 CCGAGCTGTGCTGCAAAAGGCTCTTGAAGTGTGGGCTGCGGCTGCGGCTGCGGCGAG 426
 QY 371 CCGAAGAGTCCCGAGGGGGGCTGCAAGCTGAGCTGAGCGCTGCGGAGACTGGCCAGAG 430
 Db 427 CCGAAGAGTCCCGAGGGGGGCTGCAAGCTGAGCTGAGCGCTGAGCGCTGCGGAGACTGGCCAGAG 486
 QY 431 CCGGCGCGCAGACCTTTCTCATTCACAGGCTGCGGCGCTTTAGTACTCAGAGCGAG 490
 Db 487 CCGGCGCGCAGACCTTTCTCATTCACAGGCTGCGGCGCTTTAGTACTCAGAGCGAG 546
 QY 491 CCGGAGAGTAAACAGGCTGCAAGCGCTTCTTCAAGTGTGCTGAGTGTGGAAGTGGGAGCC 550
 Db 547 CCGGAGAGTAAACAGGCTGCAAGCGCTTCTTCAAGTGTGCTGAGTGTGGAAGTGGGAGCC 606
 QY 551 GACGCGCGCAGACAGCGGCGAGCGCTGAGAGAGAGCGAGCGGAGCGCGGAGGCG 610
 Db 607 GACGCGCGCAGACAGCGGCGAGCGCTGAGAGAGAGCGAGCGGAGCGCGGAGGCG 666
 QY 611 GGAAGTGCAGCGGCGGAGAGCGGCGGAGGCTTGGCGGAGGAGAGTGTGCGCCAGAGGT 670
 Db 667 GGAAGTGCAGCGGCGGAGAGCGGCGGAGGCTTGGCGGAGGAGAGTGTGCGCCAGAGGT 726
 QY 671 GGAGAGCGCGCGCCCTCTGCTCAGCTGGAAGCACTGTGAGCTCTCTCCCGCTGAG 730
 Db 727 GGAGAGCGCGCGCCCTCTGCTCAGCTGGAAGCACTGTGAGCTCTCTCCCGCTGAG 786
 QY 731 CCGAGGTTTCTGTGCTGTGTTGGGCTGCGCAAGGCGGCGCTGCGCACTGCTTTGTG 790
 Db 787 CCGAGGTTTCTGTGCTGTGTTGGGCTGCGCAAGGCGGCGCTGCGCACTGCTTTGTG 846

QY	Db
79_	847
CCACACGCGCTGTGCGCGGAGACCCCGCTGACTGCTCGGAGTGTGGGCGCGCGCG	CCACACGCGCTGTGCGCGGAGACCCCGCTGACTGCTCGGAGTGTGGGCGCGCGCG
851	906
CTGTGTGTGGCGCCAGAGTTCTCGAGTGC	TGGTGTGTGGCG-CAGAGTTCTCGAGTGC
851	906
CTGTGTGTGGCGCCAGAGTTCTCGAGTGC	TGGTGTGTGGCG-CAGAGTTCTCGAGTGC

RESULT 12
LOCUS BC678639
DEFINITION BC678639 903 bp mRNA linear EST 15-JUL-2002
AGENCOURT 8192626 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6251109
VERSION 5', mRNA sequence.
BC678639
KEYWORDS BC678639.1 GI:21791330
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
cshg@nih.gov

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: L1CM2421 row: e column: 14
 High quality sequence stop: 689.
 Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6261103"
/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="N1H_MGC_112"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
N1H_MGC Library."

```

Query Match	32.7%	Score 786.8	DB 13	Length 903
Beet Local Similarity	99.1%	Pred. No. 9.2e-124		
Matches 791	Conservative 0	Mismatches 7	Indels 0	Gaps 0
QY	1503	TCACCTACACAGACACGGGGCGCTGTGGGGCGTGCTTCTGTGCTTACACATACATCT	1562	
DB	1	TCACCTACACAGACACGGGGCGCTGTGGGGCGCTTCTGTGCTTACACATACATCT	60	
QY	1563	TCCCCCTTCCTTGATTCGCTATGATGTCACCCACAGAGAGCAATTCGGACCCCCCAGG	1622	
DB	61	TCCCCCTTCCTTGATTCGCTATGATGTCACCCACAGAGAGCAATTCGGACCCCCCAGG	120	
QY	1623	GGCATCTGATGGCCATCTCCAGGTGAGCCAGGGCTCTGTGGCCCCGGTAAGCCAGC	1682	
DB	121	GGCATCTGATGGCCATCTCCAGGTGAGCCAGGGCTCTGTGTGGCCCCGGTAAGCCAGC	180	
QY	1683	AGTCCCATTCCTGGAGCTATGCTGCGGGCCAGACCTGACCCAGGGGAAGTTGCTAAGG	1742	

Db	182	AGTCCCAATCTCTGGGATCTGTGCGGGCCAGAGCTGGCCCGAGGGGAATGTCTAAAGG	240
Oy	1743	ATCTCTCCGGCTCTGGGATGTGTTCTTTTCAACACTGGGACCTGTGTCTGTGGATGAC	180
Db	241	ATCTCTTCGGGCTCTGGGATGTGTTCTTTCAACACTGGGACCTGTGTGTGGATGACC	300
Oy	1803	AAGTTTCTCCGCTCTCATATCGTACTGGAGACCTTCAGGTGAAGGGGGAGATG	186
Db	301	AAGTTTCTCCGCTCTCATATCGTACTGGAGACCTTCAGGTGAAGGGGGAGATG	360
Oy	1863	TGGCCAAACCGAGGTGGCGAAGGTCTTCGAGGCCCTAGATTTTCTTCAGAGGTGAAG	192
Db	361	TGGCCAAACCGAGGTGGCGAAGGTCTTCGAGGCCCTAGATTTTCTTCAGAGGTGAAG	420
Oy	1923	TCTATGGAATCTCTGTGCGCAGGCGATGAAAGCAGGCTGGAATGGAGCCCTAGTTCTC	198
Db	421	TCTATGGAATCTGTGCGCAGGCGATGAAAGCAGGCTGGAATGGAGCCCTAGTTCTC	480
Oy	1983	GTCGCCCCACGCTTGGAGCTTATGCAAGCTCTACACCAACGTGTCTGAAGATTGGCAG	204
Db	481	GTCGCCCCACGCTTGGAGCTTATGCAAGCTCTACACCAACGTGTCTGAAGATTGGCAG	540
Oy	2043	CTTATGCCCGGCCCGATTCTCAGAGCTCCAGAGGTCTTTGGGCCACCAAGAGACCTTCA	210
Db	541	TTTATGCCCGGCCCGATTCTCAGAGCTCCAGAGGTCTTTGGGCCACCAAGAGACCTTCA	600
Oy	2103	AACAGAGAAAGTTCGGATGGCAATGAGGAGCTTCGAGCCGACCAACCTGTCTGACCCAC	216
Db	601	AACAGAGAAAGTTCGGATGGCAATGAGGAGCTTCGAGCCGACCAACCTGTCTGACCCAC	660
Oy	2163	TGTACGTTTGGACCAAGGCTGTAGGTGGCTACCTGCCCCCTCAACACTGCCCGGTACAGG	222
Db	661	TGTACGTTTGGACCAAGGCTGTAGGTGGCTACCTGCCCCCTCAACACTGCCCGGTACAGG	720
Oy	2223	CCCTCTGGCAGAAACCTTGCAACTGTAGAACTTCCACACCTGAGGCAACTGAGAGAG	228
Db	721	CCCTCTGGCAGAAACCTTGCAACTGTAGAACTTCCACACCTGAGGCAACTGAGAGAG	780
Oy	2283	AACCTGTGGGGTGGGGG 2300	
Db	781	AACCTGTGGGGTGGGGG 798	

RESULT 13
 AL529217 LOCUS
 DEFINITION Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 AL529217
 ACCESSION CDNA clone CS0DD002YE03 5-PRIME, mRNA sequence.
 AL529217
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 873)
 Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 13, 2001 this sequence version replaced gi:12792710.
 COMMENT
 JOURNAL
 Title
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 661.f For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cg-bln/cluster.cgi?seq=CS0DD002AC020P1&cluster=661.f. Contact :
 Feng Liang Email : fliang@life tech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0DD002AC020P1.
 Location/Qualifiers

Wed Jul 7 12:07:09 2004

us-10-030-226-1.rst

Page 15

Db 847 AG 848

Search completed: July 3, 2004, 03:46:26
Job time : 617 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 2, 2004, 20:13:10 ; Search time 939 Seconds
(without alignments)
10880.635 Million cell updates/sec

Title: US-10-030-226-1
Perfect score: 2405
Sequence: 1 gcacccctcccggttctctg.....ccagctgctctgacctac 2405

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2405	100.0	2405	4 AAF27417	AAf27417 Human fat
2	2405	100.0	2430	7 ADB84114	ADb84114 Human POM
3	2405	100.0	2394	7 AAD49501	AAd49501 Human TRI
4	2392	99.5	2402	5 AAF89054	AAf89054 Human FAT
5	2387	99.3	2574	2 AAZ33977	AAz33977 Human PRO
6	2387	99.3	2574	3 AAC78481	AAc78481 Human PRO
7	2387	99.3	2574	3 AAC58239	AAc58239 Human PRO
8	2387	99.3	2574	7 ACD42510	ACd42510 Novel hum
9	2387	99.3	2574	7 ACA63545	ACa63545 Novel hum
10	2387	99.3	2574	7 ACA71709	ACa71709 Human sec
11	2387	99.3	2574	7 ABX92349	ABx92349 CDNA enco
12	2387	99.3	2574	7 ACA66090	ACa66090 Human CDN
13	2387	99.3	2574	8 ADA24640	ADa24640 Novel hum
14	2387	99.3	2574	8 ACD29691	ACd29691 Novel hum
15	2387	99.3	2574	8 ADA12301	ADa12301 Human CDN
16	2387	99.3	2574	8 ACD29106	ACd29106 Novel hum
17	2387	99.3	2574	8 ADB73607	ADb73607 Human PRO
18	2387	99.3	2574	9 ADB76323	ADb76323 Human PRO
19	2387	99.3	2574	9 ADC43749	ADc43749 Human CDN
20	2387	99.3	2574	9 ADC61509	ADc61509 Human CDN
21	2387	99.3	2574	9 ADC63473	ADc63473 Human CDN
22	2387	99.3	2574	9 ADC66573	ADc66573 Human CDN
23	2387	99.3	2574	9 ADC68697	ADc68697 Human CDN

24	2387	99.3	2574	9 ADC62757	ADc62757 Human CDN
25	2387	99.3	2574	9 ADC67822	ADc67822 Human CDN
26	2387	99.3	2574	9 ADC41142	ADc41142 Human CDN
27	2387	99.3	2574	9 ADC67197	ADc67197 Human CDN
28	2387	99.3	2574	9 ADC62133	ADc62133 Human CDN
29	2387	99.3	2574	9 ADC41766	ADc41766 Human CDN
30	2387	99.3	2574	9 ADE49135	ADe49135 Human CDN
31	2387	99.3	2574	9 ADE35189	ADe35189 Human CDN
32	2387	99.3	2574	9 ADE16303	ADe16303 Human CDN
33	2387	99.3	2574	9 ADD72918	ADd72918 Human CDN
34	2387	99.3	2574	9 ADB72276	ADb72276 Human CDN
35	2387	99.3	2574	9 ADE16927	ADe16927 Human CDN
36	2387	99.3	2574	9 ADE48435	ADe48435 Human CDN
37	2387	99.3	2574	10 ADE89536	ADe89536 Human CDN
38	2348.4	97.6	2447	4 AAH19186	AaH19186 Human sec
39	2310.8	96.1	2314	4 AAH17621	AaH17621 Human sec
40	2266.2	94.2	2319	4 AAH19223	AaH19223 Human sec
41	2119.4	88.1	2166	2 AAZ00379	AaZ00379 Nucleotid
42	2119.4	88.1	2166	5 AAF89041	AAf89041 Human FAT
43	2077	86.4	2835	7 AAD49515	AAd49515 Human TRI
44	2052.2	85.3	2339	5 AAD07655	AAd07655 Human sec
45	2052.2	85.3	2339	7 ADA39757	ADa39757 Human sec

ALIGNMENTS

RESULT 1					
AAf27417	ID	AAf27417 standard; CDNA; 2405 BP.			
XX	AC	AAf27417;			
XX	DT	24-APR-2001 (first entry)			
XX	DE	Human fatty acid transporter PSEC67 CDNA.			
XX	KW	Human, fatty acid transporter; PSEC67; long-chain fatty acid uptake;			
XX	KM	oleic acid; drug screening; gene therapy; metabolic disorder;			
XX	KW	cardiomyopathy; skeletal muscle disorders; renal failure; ss.			
XX	OS	Homo sapiens.			
XX	XX				
XX	PN	W0200104301-A1.			
XX	PD	18-JAN-2001.			
XX	PF	07-JUL-2000; 2000MO-JP004549.			
XX	PR	08-JUL-1999; 99JP-00194179.			
XX	PR	18-OCT-1999; 99US-0159586P.			
XX	PR	23-APR-2000; 2000UP-00128993.			
XX	PA	(HELI-) HELIX RES INST.			
XX	PI	MoriKawa N, Masuno Y, Ota T, Isogai T, NishiKawa T, Kawai Y;			
XX	PI	WPI; 2001-138349/14.			
XX	DR	P-Psds; AAB60388.			
XX	PT	Fatty acid transporter protein and encoded gene PSEC67 cloned from human			
XX	PT	CDNA library, with activity of oleic acid incorporation, useful as target			
XX	PT	molecule of preventives or remedies of fatty-acid metabolic disorders.			
XX	PS	Claim 1; Page 44-48; 58pp; Japanese.			
XX	CC	The invention relates to a novel human fatty acid transporter, PSEC67			
XX	CC	(AAB60388), and to CDNA encoding it (AAf27417). PSEC67 is responsible for			
XX	CC	the uptake of oleic acid into cells. The invention also relates to			
XX	CC	vectors and host cells comprising a PSEC67 nucleic acid; the recombinant			
XX	CC	production of PSEC67; an antibody against PSEC67; methods of screening			
XX	CC	for compounds which can regulate the uptake of long-chain fatty acids			
XX	CC	into cells; and the compounds thus identified. The PSEC67 protein and the			

Db 1981 GCGTCCCCCCAGCTTTGACCTTATGACGCTTACACCAACGCTGTGAGAACTTGCC 2040
Qy 2041 ACCTTATGCCCCCGCCGATTCCTAGGCTCCAGAGTCTTTGGCCACCAAGAGACTT 2100
Db 2041 ACCTTATGCCCCCGCCGATTCCTAGGCTCCAGAGTCTTTGGCCACCAAGAGACTT 2100
Qy 2101 CAACAGAGAGAAAGTTCGATGAGCAATGAGAGGCTTCGACCCAGCACTCTGTGACCC 2160
Db 2101 CAACAGAGAGAAAGTTCGATGAGCAATGAGAGGCTTCGACCCAGCACTCTGTGACCC 2160
Qy 2161 ACTGACGTTTGTGACCAAGCTGTAGTGTCTTACCTGCCCCCTCACAACTGCCCCGATACG 2220
Db 2161 ACTGACGTTTGTGACCAAGCTGTAGTGTCTTACCTGCCCCCTCACAACTGCCCCGATACG 2220
Qy 2221 CGCCCTCTGGGAGAGAAACCTTCGAATGAGAACTTCACACCTGAGGACCTGAGAGA 2280
Db 2221 CGCCCTCTGGGAGAGAAACCTTCGAATGAGAACTTCACACCTGAGGACCTGAGAGA 2280
Qy 2281 GGAACCTCTGTGGGAGTGGGGGCGGTGACAGGTGTACTGTGGGCTGTGAGGATCTTTTCTATA 2340
Db 2281 GGAACCTCTGTGGGAGTGGGGGCGGTGACAGGTGTACTGTGGGCTGTGAGGATCTTTTCTATA 2340
Qy 2341 CGAAGACTGCGGTCTACTATTTTGTATATAAATGTGCTGAGACTGATCCAGCTGTCTCTGA 2400
Db 2341 CGAAGACTGCGGTCTACTATTTTGTATATAAATGTGCTGAGACTGATCCAGCTGTCTCTGA 2400
Qy 2401 CCTAC 2405
Db 2401 CCTAC 2405

RESULT 2
ADA84114
ID ADA84114 standard; DNA; 2430 BP.
XX
AC ADA84114;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human POM124 gene.
XX
XX human; marker: expressed sequence tag; EST; arabidopsis; tumour;
XX stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;
XX vaccine; ds; gene.
XX
OS Homo sapiens.
XX
PN MO2002103028-A2.
XX
PD 27-DEC-2002.
XX
PF 30-MAY-2002; 2002MO-IB004189.
XX
XX 30-MAY-2001; 2001US-0293999P.
XX 22-OCT-2001; 2001US-0330457P.
XX 19-FEB-2002; 2002US-0357144P.
XX
PA (BIOM-) BIOMEDICAL CENT.
XX
PI Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;
XX WPI; 2003-175241/17.
XX DR P-PSDB; ADA84115.
XX
XX Determining if a nucleic acid is a marker for a phenotype/cell type of
XX interest, by global comparison of expressed sequence tags known to be
XX expressed in the phenotype/cell type with all ESTs expressed in normal
XX tissue.
XX
XX Claim 23; Page 490-491; 516pp; English.
XX
XX The invention relates to a novel method for determining if a nucleic acid
XX is a marker for a predetermined phenotype/cell type of interest from a

CC biological species. The method comprises performing a global comparison
CC of a group of expressed sequence tags (ESTs) known to be expressed in the
CC phenotype/cell type of interest with all ESTs expressed in normal tissue
CC in order to identify ESTs that are preferentially expressed in the
CC phenotype/cell type of interest. A method of the invention is useful for
CC determining whether a nucleic acid is a marker for a predetermined
CC phenotype or cell type of interest from a biological species, preferably
CC Arabidopsis or human. The cell type of interest is an abnormal cell such
CC as a tumour cell, and the predetermined phenotype is a stress-induced
CC phenotype such as hyperosmotic stress or high salt conditions. A method
CC of the invention is also useful for determining the progression of colon
CC cancer in a human, for detecting a tumour cell, and for regulating or
CC preventing the growth of a tumour cell. An antibody of the invention is
CC useful for detecting the absence or presence of peptides encoded by
CC tumour-associated markers. A polypeptide of the invention is useful as an
CC immunogen for vaccinating an animal. The present sequence encodes a
CC tumour-associated antigen of the invention.
XX
SQ Sequence 2430 BP; 430 A; 740 C; 781 G; 479 T; 0 U; 0 Other;
Query Match 100.0%; Score 2405; DB 7; Length 2430;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCACTCTCTCCGCGGTTTCTGCTCTCCGCGGTGTGAGTGTGGGGGCTTGGGTGGGAAAT 60
Db 8 GCACTCTCTCCGCGGTTTCTGCTCTCCGCGGTGTGAGTGTGGGGGCTTGGGTGGGAAAT 67
Qy 61 GGGGAGTGTGCGACGACGCGCGCTCCCTGGAAGAGAAAGTCTCAGCTAAGACGACGGC 120
Db 68 GGGGAGTGTGCGACGACGCGCGCTCCCTGGAAGAGAAAGTCTCAGCTAAGACGACGGC 127
Qy 121 CCTAGTTTGGAGAGGAGATCAGGAGATGTTTTCGAGCGGCTGGAACCAACGCTGCC 180
Db 128 CCTAGTTTGGAGAGGAGATCAGGAGATGTTTTCGAGCGGCTGGAACCAACGCTGCC 187
Qy 181 GATAGAGAGAGCGGCGCTCCATGCGTCCCTCGCTGCTGCGCTCGCTGCTGTGTGATAC 240
Db 188 GATAGAGAGAGCGGCGCTCCATGCGTCCCTCGCTGCTGCGCTCGCTGCTGTGTGATAC 247
Qy 241 GCTGCTGCTGCTGAAGCTTACCTCTGCGCGAGTTGCGCTGCTTCCGCGGACTTGGC 300
Db 248 GCTGCTGCTGCTGAAGCTTACCTCTGCGCGAGTTGCGCTGCTTCCGCGGACTTGGC 307
Qy 301 CTTTGCCTGCGAGCTTGTGTGCGAAAGAGGCTTTTCAAGCTCGGCGCTTGGCGCGCGC 360
Db 308 CTTTGCCTGCGAGCTTGTGTGCGAAAGAGGCTTTTCAAGCTCGGCGCTTGGCGCGCGC 367
Qy 361 TGCCTGCGAGCGCGGAGAGTCCGAGGAGGAGGCTGCACTGCGCTGCGCTCGCGAGACT 420
Db 368 TGCCTGCGAGCGCGGAGAGTCCGAGGAGGAGGCTGCACTGCGCTGCGCTCGCGAGACT 427
Qy 421 GGCCTGAGAGCGCGCGCGCGCAACCTTTCTCATTCACGGCTGCGCGGCTTTTACTATC 480
Db 428 GGCCTGAGAGCGCGCGCGCGCAACCTTTCTCATTCACGGCTGCGCGGCTTTTACTATC 487
Qy 481 AGAGCGGAGCGCGAGAGTAAGAGGAGGCTGCAAGCGCTTCTCAAGTGTGCGCTAGAGTGGGA 540
Db 488 AGAGCGGAGCGCGAGAGTAAGAGGAGGCTGCAAGCGCTTCTCAAGTGTGCGCTAGAGTGGGA 547
Qy 541 CTGGGAGCCCGACCGCGCGCGACGCGCGAGGAGGAGGCTGAGAAAGCGAGCGCGACG 600
Db 548 CTGGGAGCCCGACCGCGCGCGACGCGCGAGGAGGAGGCTGAGAAAGCGAGCGCGACG 607
Qy 601 GCGGAGCGCGAGAGTACAGCGCGCGCGAGAGCGCGCGAGTTCGCGAGAGGAGCGAGTGC 660
Db 608 GCGGAGCGCGAGAGTACAGCGCGCGCGAGAGCGCGCGAGTTCGCGAGAGGAGCGAGTGC 667
Qy 661 GCGGAGAGTGTGAGAGAGCGCGCGCTCTGTACCTGTGAGCAACTGTGAGCTGTCTCT 720
Db 668 GCGGAGAGTGTGAGAGAGCGCGCGCTCTGTACCTGTGAGCAACTGTGAGCTGTCTCT 727
Qy 721 CCGCGCTGCGCGCGAGATTTCTGTGTGCTGTGCTTCCGCGCTGCGCAAGCGCGCTTGGCAC 780

Dp	728	CCCGGCTGGCCCAAGTTTCTGTGGCTTGGGCTGGCCCAAGGCGGCTTCGGCAC	787
Qy	781	TGCGTTTGTGCCACCCCGCTGGCGCCGGGGCCCCCTGTGTCACATGCTCCGAGCTGGG	840
Dp	788	TGCCTTTGTGCCACCGCCCTGGCGCCGGGGCCCCCTGTGTCACATGCTCCGAGCTGGG	847
Qy	841	CGCGCGCGCTGTGTGTGGCGCCAGATTTCTGTGAGTCCCTGTGAGCCGGAACCTGCCGC	900
Dp	848	CGCGCGCGCTGTGTGTGGCGCCAGATTTCTGTGAGTCCCTGTGAGCCGGAACCTGCCGC	907
Qy	901	CCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGGCCAGGAACCCACCTGTGGAAT	960
Dp	908	CCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGGCCAGGAACCCACCTGTGGAAT	967
Qy	961	TAGCGATTGTGTGCTGAAGTGTCCGTGAATGTGAATGGGCCAGTGTCCAGATCTCTC	1020
Dp	968	TAGCGATTGTGTGCTGAAGTGTCCGTGAATGTGAATGGGCCAGTGTCCAGATCTCTC	1027
Qy	1021	TTCCCGCCAGAGCATACAGACAGTGGCTGTACATCTTCACTCTGGGACACAGGGCT	1080
Dp	1028	TTCCCGCCAGAGCATACAGACAGTGGCTGTACATCTTCACTCTGGGACACAGGGCT	1087
Qy	1081	CCCCAAGGCTGTGGATCGATCATCTGAAGATCTGCAATGCGAGGGCTTCTATCAGCT	1140
Dp	1088	CCCCAAGGCTGTGGATCGATCATCTGAAGATCTGCAATGCGAGGGCTTCTATCAGCT	1147
Qy	1141	GTGTGTGTCCACAGAGAAATGTGAATCACTCGGCCCTCCACCTTACCAATGTCCGG	1200
Dp	1148	GTGTGTGTCCACAGAGAAATGTGAATCACTCGGCCCTCCACCTTACCAATGTCCGG	1207
Qy	1201	TTCCCTGCTGGGACCTGTGGGCTGTGATGGGCAATTTGGGGCCAGAGTGTCTGAAATCCAA	1260
Dp	1208	TTCCCTGCTGGGACCTGTGGGCTGTGATGGGCAATTTGGGGCCAGAGTGTCTGAAATCCAA	1267
Qy	1261	GTTTCTGGGCTGTGTGTGTGGGAAATTTGCCAGACAGCAGGGTGTCTCACTA	1320
Dp	1268	GTTTCTGGGCTGTGTGTGTGGGAAATTTGCCAGACAGCAGGGTGTCTCACTA	1327
Qy	1321	CATTGGGAGACTGTGCGATACCTTGTCAACAGCCCCCGAGCAAGGCAAGACGTGGCCA	1380
Dp	1328	CATTGGGAGACTGTGCGATACCTTGTCAACAGCCCCCGAGCAAGGCAAGACGTGGCCA	1387
Qy	1381	TAAAGTCCGAGCTGGCAGTGGGCGCGGCTGGCCCAATACCTTGGAGCCGTTTGTGCG	1440
Dp	1388	TAAAGTCCGAGCTGGCAGTGGGCGCGGCTGGCCCAATACCTTGGAGCCGTTTGTGCG	1447
Qy	1441	GCGCTTGGGCGCCCTGTGAGGTGTGTGAGACATATGGACGTACAGAGGGCAACGTGGCAC	1500
Dp	1448	GCGCTTGGGCGCCCTGTGAGGTGTGTGAGACATATGGACGTACAGAGGGCAACGTGGCAC	1507
Qy	1501	CATCAACTACACAGACAGCGGGGCGCTGTGGGCGCTGTGCTTCTGCTTTCACACATAT	1560
Dp	1508	CATCAACTACACAGACAGCGGGGCGCTGTGGGCGCTGTGCTTCTGCTTTCACACATAT	1567
Qy	1561	CTTCCCTTCTCTTATTCGTATGTGATGTCAACACAGAGGCCAATTTGGGACCCCA	1620
Dp	1568	CTTCCCTTCTCTTATTCGTATGTGATGTCAACACAGAGGCCAATTTGGGACCCCA	1627
Qy	1621	GGGCACTGTATGACACATCTTCAGGTAGGACAGGGCTGTGTGGCCCGGTAAGCCA	1680
Dp	1628	GGGCACTGTATGACACATCTTCAGGTAGGACAGGGCTGTGTGGCCCGGTAAGCCA	1687
Qy	1681	GCAAGTCCCAATTTCTGGGCTGTATGTGTGGGCGCAAGAGTGTGCCAGGGGAAATTTCTAA	1740
Dp	1688	GCAAGTCCCAATTTCTGGGCTGTATGTGTGGGCGCAAGAGTGTGCCAGGGGAAATTTCTAA	1747
Qy	1741	GGATGTCTTCCGAGCTGGGATGTTTCTTCAACACTGGGAGCTGTGTGTCTGGATGA	1800
Dp	1748	GGATGTCTTCCGAGCTGGGATGTTTCTTCAACACTGGGAGCTGTGTGTCTGGATGA	1807
Qy	1801	CCAAGGTTTCTCGCTTCCATGATCTGAGACACCTTCAGTGTGAAAGGGGAGAA	1860

Dp	1808	CCAAGGTTTCTCGCTTCCATGATCTGAGACACCTTCAGTGTGAAAGGGGAGAA	1867
Qy	1861	TGTGGCCACACCGAGGTGGCAGAGGTCTTGCAGGCCCTAGATTTTCTTCCAGAGGTGA	1920
Dp	1868	TGTGGCCACACCGAGGTGGCAGAGGTCTTGCAGGCCCTAGATTTTCTTCCAGAGGTGA	1927
Qy	1921	CGCTTATGAGATCACTGTGCCAGGSCATGAAGGCGGGCTGGAATGGAGCCCTAGTCT	1980
Dp	1928	CGCTTATGAGATCACTGTGCCAGGSCATGAAGGCGGGCTGGAATGGAGCCCTAGTCT	1987
Qy	1981	GGCTCCCCCCCAAGCTTTTGAACCTTATGACGCTTACACCCACGCTGTGAACTTGCC	2040
Dp	1988	GGCTCCCCCCCAAGCTTTTGAACCTTATGACGCTTACACCCACGCTGTGAACTTGCC	2047
Qy	2041	ACCTTATGCCCCGCCCTCCATTTCTTAGGCTCCAGAGATCTTTGGCCACACAGAGACTT	2100
Dp	2048	ACCTTATGCCCCGCCCTCCATTTCTTAGGCTCCAGAGATCTTTGGCCACACAGAGACTT	2107
Qy	2101	CAACACAGAGAAATTTGGATGGCAGAAATGAGGGCTTGCAGCCACAGACCTGTGTACC	2160
Dp	2108	CAACACAGAGAAATTTGGATGGCAGAAATGAGGGCTTGCAGCCACAGACCTGTGTACC	2167
Qy	2161	ACTGTAGTTCTTGACACAGGCTGTATGTGTCTTACTGTCCCTCACTGCTGCGGTACAG	2220
Dp	2168	ACTGTAGTTCTTGACACAGGCTGTATGTGTCTTACTGTCCCTCACTGCTGCGGTACAG	2227
Qy	2221	CGCCCTCTGGAGAGAACTTGCATCTGAGAACTTCAACCTGACCTGAGGACCTGAGAGA	2280
Dp	2228	CGCCCTCTGGAGAGAACTTGCATCTGAGAACTTCAACCTGACCTGAGGACCTGAGAGA	2287
Qy	2281	GGAACCTGTGTGGGCTGTGGGCGCGTGTGAGGTGTACTGGCTGTGAGGATCTTTCTATA	2340
Dp	2288	GGAACCTGTGTGGGCTGTGGGCGCGTGTGAGGTGTACTGGCTGTGAGGATCTTTCTATA	2347
Qy	2341	CCAGAACTGGGCTCACTTTTGTATTAATGATGGCTGGAGCTGATCAGCTGTCTGTA	2400
Dp	2348	CCAGAACTGGGCTCACTTTTGTATTAATGATGGCTGGAGCTGATCAGCTGTCTGTA	2407
Qy	2401	CCTAC 2405	
Dp	2408	CCTAC 2412	

RESULT 3

AAD49501

ID AAD49501 standard; cDNA; 2994 BP.

XX AAD49501;

AC 24-VR-2003 (first entry)

DT XX

DE Human TRICH-3 cDNA.

XX Human; transporter and ion channel; TRICH; atherosclerosis; cancer;

KM Gene therapy; gene; ss.

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 50..2485

FT *tag= a

FT /product= "Human TRICH protein"

XX

WO200283712-A2.

XX

PD 24-OCT-2002.

PF 12-APR-2002; 2002MWO-US011760.

XX

PR 12-APR-2001; 2001US--0283440P.

PR 20-APR-2001; 2001US-0285592P.

PR 27-APR-2001; 2001US-0287263P.

PR 04-MAY-2001; 2001US-0288666P.

Db 1855 GGGGCACTGATGAGCCACATCTCCAGTGTAGCCAGGAGCTGTGTGCCCCCGTAAAGCA 1914
 Qy 1681 GCAGTCCCATCTCTGGGCTATAGCTGGGGGCGCAGAGCTGGGGGCAAGGGAATCTCTAA 1740
 Db 1915 GCAGTCCCATCTCTGGGCTATAGCTGGGGGCGCAGAGCTGGGGGCAAGGGAATCTCTAA 1974
 Qy 1741 GGAATCTCTCTGGGCTATAGCTGGGGGCGCAGAGCTGGGGGCAAGGGAATCTCTAA 1800
 Db 1975 GGAATCTCTCTGGGCTATAGCTGGGGGCGCAGAGCTGGGGGCAAGGGAATCTCTAA 2034
 Qy 1801 CCAAGGTTTCTCTGGGCTATAGCTGGGGGCGCAGAGCTGGGGGCAAGGGAATCTCTAA 1860
 Db 2035 CCAAGGTTTCTCTGGGCTATAGCTGGGGGCGCAGAGCTGGGGGCAAGGGAATCTCTAA 2094
 Qy 1861 TGTGCGCACAAACCGAGAGTGGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 1920
 Db 2095 TGTGCGCACAAACCGAGAGTGGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2154
 Qy 1921 CGCTATGAGAGTCTCTGAGAGTGGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 1980
 Db 2155 CGCTATGAGAGTCTCTGAGAGTGGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2214
 Qy 1981 GCGTCCCGCCCAAGCTTGTGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2040
 Db 2215 GCGTCCCGCCCAAGCTTGTGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2274
 Qy 2041 ACCTTATGCGCGCGCGCGCTTCTTCAAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2100
 Db 2275 ACCTTATGCGCGCGCGCGCTTCTTCAAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2334
 Qy 2101 CAAAGCAGAAAGTGTGAGAGTGGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2160
 Db 2335 CAAAGCAGAAAGTGTGAGAGTGGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2394
 Qy 2161 ACTGATGCTTGTGAGAGTGGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2220
 Db 2395 ACTGATGCTTGTGAGAGTGGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2454
 Qy 2221 CGCGCCCGCGCGCGCGCTTCTTCAAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2280
 Db 2455 CGCGCCCGCGCGCGCGCTTCTTCAAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2514
 Qy 2281 GGAAGTCTGTGGGCGCGCGCTTGTGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2340
 Db 2515 GGAAGTCTGTGGGCGCGCGCTTGTGAGAGTCTTCCAGAGGCTTGTCTTCTTCAAGAGTGA 2574
 Qy 2341 CCAAGACTCGGCTCACTATTTTGTAAATAATGTGCTGAGAGTCTTCCAGAGTCTTGA 2400
 Db 2575 CCAAGACTCGGCTCACTATTTTGTAAATAATGTGCTGAGAGTCTTCCAGAGTCTTGA 2634
 Qy 2401 CCTAC 2405
 Db 2635 CCTAC 2639
 RESULT 4
 AAF89054 standard; DNA; 2402 BP.
 XX AAF89054;
 AC AAF89054;
 DT 06-JUL-2001 (first entry)
 XX
 DE Human FATP3 coding sequence SEQ ID NO: 116.
 XX
 KW Fatty acid transport protein; FATP; human; mouse; rat; rice blast fungus;
 XX Yeast; fat absorption; obesity; diabetes; heart disease; hyperlipidemia;
 XX weight control; tuberculosis; TB; anti-fungal; ds.
 OS Homo sapiens.
 XX
 PN WO200121795-A2.
 XX

PD 29-MAR-2001.
 XX
 PF 21-SEP-2000; 2000WO-US025891.
 XX
 PR 23-SEP-1999; 99US-00405504.
 PR 16-DEC-1999; 99US-00465280.
 PR 17-FEB-2000; 2000US-00506252.
 PR 06-JUL-2000; 2000US-00611197.
 XX
 PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Stahl A, Hirsch DJ, Lodish HF, Gimeno RE, Tartaglia LA;
 XX
 DR WPI; 2001-354783/37.
 DR P-PSDB; AAB83279.
 XX
 FT New fatty acid transport proteins (FATPs) useful for the manufacture of
 FT medicament for treating obesity, diabetes and heart disease.
 XX
 PS Claim 1; Fig 111; 287pp; English.
 XX
 CC The present invention provides the protein and coding sequences of fatty
 CC acid transport proteins (FATPs) from a number of species, including
 CC FATP1, FATP2, FATP3, FATP4, FATP5 and FATP6 from the human, FATP1-FATP5
 CC from the mouse, FATP4 and b from C. elegans, the grisea, Mycobacterium
 CC nidulans, Drosophila, zebrafish, Magnaporthe grisea, Mycobacterium
 CC tuberculosis and Cochliobolus heterostrophus. The FATP from M.
 CC treat TB. That from M. grisea (also known as rice blast fungus) can be
 CC used to develop anti-fungal agents capable of preventing infection of
 CC rice. Those from the human can be used to develop treatments for
 CC diabetes, heart disease, obesity, hyperlipidemia and weight control. The
 CC present sequence is one of the sequences described in the exemplification
 CC of the invention
 CC
 SQ Sequence 2402 BP; 418 A; 728 C; 780 G; 476 T; 0 U; 0 Other;
 Query Match 99.5%; Score 2392; DB 5; Length 2402;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 9 CCGGGGTTTCTGCTCTCCGCGCGCTGTGAGAGTGGGCGCGCTGTGAGAGTGGGCGTGT 68
 Db 1 CCGGGGTTTCTGCTCTCCGCGCGCTGTGAGAGTGGGCGCGCTGTGAGAGTGGGCGTGT 60
 Qy 69 GCCAGCGCACGCGCGCTCTCTGGAAGAGAGTCTCACTAAGAGAGGCGGCTTAAAGTT 128
 Db 61 GCCAGCGCACGCGCGCTCTCTGGAAGAGAGTCTCACTAAGAGAGGCGGCTTAAAGTT 120
 Qy 129 TTCGGAAGGAGGATCAGGAGTGTTCGAGCGGCTGTGAACGAGAGCGGTGAGAG 188
 Db 121 TTCGGAAGGAGGATCAGGAGTGTTCGAGCGGCTGTGAACGAGAGCGGTGAGAG 180
 Qy 189 AAGCGGCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248
 Db 181 AAGCGGCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 Qy 249 TGTGAGAGTCAACCTCTGCGCGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 308
 Db 241 TGTGAGAGTCAACCTCTGCGCGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 300
 Qy 309 TGTGAGTCTGTGCTCAAAAGGCTTTCGAGCTGCGCGCTGCGCGCGCTGCGCGCG 368
 Db 301 TGTGAGTCTGTGCTCAAAAGGCTTTCGAGCTGCGCGCGCTGCGCGCGCTGCGCGCG 360
 Qy 369 ACCCGGAGAGTCCCGAGGGGGGCTGTGAGCTGCGCGCGCTGCGCGCGCTGCGCGCG 428
 Db 361 ACCCGGAGAGTCCCGAGGGGGGCTGTGAGCTGCGCGCGCTGCGCGCGCTGCGCGCG 420
 Qy 429 AGCGCGCGCGCACACTTCTCATTCAGAGGCTGCGCGCGCTTGACTAGAGGCGG 488

Db 421 AGCGCGCGGACACCTTTCTCATTTACGGCTCGCGCGCTTTAGACTACAGAGCGG 480
 QY 489 AGCGGAGATTAACAGGGCTGCACGCGCTTCTTACGTGCGTAGGCTGGAGCTGGGAC 548
 Db 481 AGCGGAGATTAACAGGGCTGCACGCGCTTCTTACGTGCGTAGGCTGGAGCTGGGAC 540
 QY 549 CCGACGCGCGGACAGCGCGGCGAGCGCTGAGAAAGCGAGCGGCGAGCGCGGAG 608
 Db 541 CCGACGCGCGGACAGCGCGGCGAGCGCTGAGAAAGCGAGCGGCGAGCGCGGAG 600
 QY 609 CCGGAGATTAACAGGGCTGCACGCGCTTCTTACGTGCGTAGGCTGGAGCTGGGAC 668
 Db 601 CCGGAGATTAACAGGGCTGCACGCGCTTCTTACGTGCGTAGGCTGGAGCTGGGAC 660
 QY 669 GTGAGAGAGCGCGCGCGCTTCTTACGTGAGCACTGTGAGCGCTGCTCCCGCTG 728
 Db 661 GTGAGAGAGCGCGCGCGCTTCTTACGTGAGCACTGTGAGCGCTGCTCCCGCTG 720
 QY 729 GCGGAGATTTCTGTGAGCTGTGAGCGCGGCGAGCGCGGCGCTGCGCACTGCGCTT 788
 Db 721 GCGGAGATTTCTGTGAGCTGTGAGCGCGGCGAGCGCGGCGCTGCGCACTGCGCTT 780
 QY 789 TGCGGAGCGCGCGCGCGCGCGCTGCGCACTGCGCTGCGCACTGCGCACTGCGCG 848
 Db 781 TGCGGAGCGCGCGCGCGCGCGCTGCGCACTGCGCTGCGCACTGCGCACTGCGCG 840
 QY 849 CGCTGCTGTGCGCGCGCAAGATTTCTGAGTCCCTGAGACCGGAGCTGCGCGCGCG 908
 Db 841 CGCTGCTGTGCGCGCGCAAGATTTCTGAGTCCCTGAGACCGGAGCTGCGCGCGCG 900
 QY 909 CCATGAGGCTCCACCTGTGAGCGCGGCGAGAACCCACCTGCTGAGATTAAGCAT 968
 Db 901 CCATGAGGCTCCACCTGTGAGCGCGGCGAGAACCCACCTGCTGAGATTAAGCAT 960
 QY 969 TGCTGCTGTGAGCTGTGAGCGCGGCGAGAACCCACCTGCTGAGATTAAGCAT 1028
 Db 961 TGCTGCTGTGAGCTGTGAGCGCGGCGAGAACCCACCTGCTGAGATTAAGCAT 1020
 QY 1029 AGAGATTAACAGAGCGGCGCTGATCTTCACTCTGAGCACTGAGCACTGAGCACTG 1088
 Db 1021 AGAGATTAACAGAGCGGCGCTGATCTTCACTCTGAGCACTGAGCACTGAGCACTG 1080
 QY 1089 CTGCTGAGATCATCTGAGATCTGAGATCTGAGATCTGAGATCTGAGATCTGAG 1148
 Db 1081 CTGCTGAGATCATCTGAGATCTGAGATCTGAGATCTGAGATCTGAGATCTGAG 1140
 QY 1149 TCCACAGAGAGATGATCTTACCTGCGCTTCCCACTTACCACTGCTGCGCTTCC 1208
 Db 1141 TCCACAGAGAGATGATCTTACCTGCGCTTCCCACTTACCACTGCTGCGCTTCC 1200
 QY 1209 TGGGATCGTGGGCTGATGAGGCTTGGGCGCAGAGTGTGCTGAATCTGCTGCG 1268
 Db 1201 TGGGATCGTGGGCTGATGAGGCTTGGGCGCAGAGTGTGCTGAATCTGCTGCG 1260
 QY 1269 CTGCTGAGTCTGAGAAATGCTGAGCACTGAGCACTGAGCACTGAGCACTGAG 1328
 Db 1261 CTGCTGAGTCTGAGAAATGCTGAGCACTGAGCACTGAGCACTGAGCACTGAG 1320
 QY 1329 AGCTGAGCGATATCTTGTCAACAGCGCGCGGCGAGCACTGAGCACTGAGCACTGAG 1388
 Db 1321 AGCTGAGCGATATCTTGTCAACAGCGCGCGGCGAGCACTGAGCACTGAGCACTGAG 1380
 QY 1389 GCGTGGAGTGGGAGCGCGGCTGCGCCAGATACCTGAGAGCGCTTTTGTGCGCGCT 1448
 Db 1381 GCGTGGAGTGGGAGCGCGGCTGCGCCAGATACCTGAGAGCGCTTTTGTGCGCGCT 1440
 QY 1449 GCGCGCTGAGGCTGAGCACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1508
 Db 1441 GCGCGCTGAGGCTGAGCACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500
 QY 1509 ACACAGAGAGCGGCGCGCTGAGGAGCGTGTCTTCTGCTTACAGAGATATCTTCCCT 1568
 Db 1501 ACACAGAGAGCGGCGCGCTGAGGAGCGTGTCTTCTGCTTACAGAGATATCTTCCCT 1560

QY 1569 TCTCTGATGTCGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1628
 Db 1561 TCTCTGATGTCGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 QY 1629 GTATGCGCATATCCAGGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1688
 Db 1621 GTATGCGCATATCCAGGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1680
 QY 1689 CATTCGTGGGCTATGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1748
 Db 1681 CATTCGTGGGCTATGTCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1740
 QY 1749 TCGGCTGCGGATGTTTCTTCAACACTGAGGAGCTGCTGCTGCTGCTGCTGCTG 1808
 Db 1741 TCGGCTGCGGATGTTTCTTCAACACTGAGGAGCTGCTGCTGCTGCTGCTGCTG 1800
 QY 1809 TTCTCGGCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1868
 Db 1801 TTCTCGGCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 QY 1869 CAACGAGTGGGAGGTCCTGAGCGCGCTGAGCTTCTTCAAGAGTGAACGCTTATG 1928
 Db 1861 CAACGAGTGGGAGGTCCTGAGCGCGCTGAGCTTCTTCAAGAGTGAACGCTTATG 1920
 QY 1929 GAGTCACTGTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1988
 Db 1921 GAGTCACTGTCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
 QY 1989 CCCAGCTTTGAGCCTTATGACCTGATGACCCAGCTGCTGAGAACTTGCACTTATG 2048
 Db 1981 CCCAGCTTTGAGCCTTATGACCTGATGACCCAGCTGCTGAGAACTTGCACTTATG 2040
 QY 2049 CCGGCGCGGATTCCTCAAGCTTCCAGAGCTCTTGGCGCAGAGACCTTCAACAG 2108
 Db 2041 CCGGCGCGGATTCCTCAAGCTTCCAGAGCTCTTGGCGCAGAGACCTTCAACAG 2100
 QY 2109 AGAAGTGGGATGAGAAATGAGGCGCTGAGACCCAGCACTGCTGAGCACTGAG 2168
 Db 2101 AGAAGTGGGATGAGAAATGAGGCGCTGAGACCCAGCACTGCTGAGCACTGAG 2160
 QY 2169 TTCTGAGCAAGCTTGAAGTGTCTTACCTGCTTCACTGCTTCACTGCTTCACTG 2228
 Db 2161 TTCTGAGCAAGCTTGAAGTGTCTTACCTGCTTCACTGCTTCACTGCTTCACTG 2220
 QY 2229 TGGCAGAAACCTTGAATCTGAGAACTTCAACCTGAGGCACTGAGAGAGAACTCT 2288
 Db 2221 TGGCAGAAACCTTGAATCTGAGAACTTCAACCTGAGGCACTGAGAGAGAACTCT 2280
 QY 2289 GTGGGCTGGGCGCGTGGCAGGTCGAGGTCGAGGTCGAGGTCGAGGTCGAGGTC 2348
 Db 2281 GTGGGCTGGGCGCGTGGCAGGTCGAGGTCGAGGTCGAGGTCGAGGTCGAGGTC 2340
 QY 2349 GCGGTACATATTTTGTAAATGATGATGATGATGATGATGATGATGATGATGAT 2400
 Db 2341 GCGGTACATATTTTGTAAATGATGATGATGATGATGATGATGATGATGATGAT 2392

RESULT 5
 AA233977
 ID AA233977 standard; cDNA; 2574 bp.
 XX AA233977;
 AC
 XX
 DT 07-DEC-1999 (first entry)
 XX
 DE Human PRO703 nucleotide sequence.
 XX
 XX Human; PEO; EST; expressed sequence tag; PCR primer; hybridisation;
 KM probe; blood coagulation disorder; cancer; cellular adhesion disorder;
 KM secreted protein; transmembrane protein; ss.
 OS Homo sapiens.

[illegible]

PR	07-MAY-1998;	98US-0084643P.
PR	13-MAY-1998;	98US-0085323P.
PR	13-MAY-1998;	98US-0085338P.
PR	13-MAY-1998;	98US-0085339P.
PR	15-MAY-1998;	98US-0085573P.
PR	15-MAY-1998;	98US-0085579P.
PR	15-MAY-1998;	98US-0085680P.
PR	15-MAY-1998;	98US-0085682P.
PR	15-MAY-1998;	98US-0085689P.
PR	15-MAY-1998;	98US-0085697P.
PR	15-MAY-1998;	98US-0085700P.
PR	15-MAY-1998;	98US-0085704P.
PR	18-MAY-1998;	98US-0086023P.
PR	22-MAY-1998;	98US-0086392P.
PR	22-MAY-1998;	98US-0086414P.
PR	22-MAY-1998;	98US-0086430C.
PR	22-MAY-1998;	98US-0086485P.
PR	28-MAY-1998;	98US-0087098P.
PR	28-MAY-1998;	98US-0087105C.
PR	28-MAY-1998;	98US-0087209P.
PR	30-JUL-1998;	98US-0094651P.
PR	11-SEP-1998;	98US-0100038P.

PA (GERTH) GENENTECH INC.
XX
XX
PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
XX WPI; 1999-551358/46.
XX
XX P-PSDB; AAY41699.
XX
XX New secreted and transmembrane polypeptides and their polynucleotides
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders.
XX
XX Claim 2, Fig 38, 530pp, English.
XX

The present invention describes nucleotide and transmembrane polypeptides and their polynucleotides. The nucleotide sequences are useful as sources of probes, primers, for chromosome mapping, and for generation of antisense sequences. They can also be used to create transgenic animals. The proteins can be used to treat a variety of diseases and disorders, depending on their function. Diseases that may be treated include blood coagulation disorders, cancers and cellular adhesion disorders. They may also be used to raise antibodies. AA23391 to AA24338, and AA41685 to AA41774 represent polynucleotide and polypeptide sequence given in the exemplification of the present invention

Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;

Query Match 99.3%; Score 2387; DB 2; Length 2574;

Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	16	CTGCTCTCCGCCCGCTGTGGAGTGGTGGGGGCGCTGGAGTGGGAATGGCGTGTGCACGGCA	77
Db	74	CTGCTCTCCGCCCGCTGTGGAGTGGTGGGGGCGCTGGAGTGGGAATGGCGTGTGCACGGCA	133
QY	78	CGCGCGCTCCCTGGAAGAGAAAGTCTCAGCTAGAACGACGGCCCTAGTTTTTCGAAAG	137
Db	134	CGCGCGCTCCCTGGAAGAGAAAGTCTCAGCTAGAACGACGGCCCTAGTTTTTCGAAAG	193
QY	138	GAGAGATCAGAGATGTTTTGCGAGCGGCTTGGAAACCAAGCGGTGCCGATAGAGAAAGCGGGCT	197
Db	194	GAGAGATCAGAGATGTTTTGCGAGCGGCTTGGAAACCAAGCGGTGCCGATAGAGAAAGCGGGCT	253
QY	198	CCATGCGCTCCCTCCGCTGCTGCGCCCTGCTGCTGCTGTACGACTGCTGCTCTGAAG	257
Db	254	CCATGCGCTCCCTCCGCTGCTGCGCCCTGCTGCTGCTGTGTACGACTGCTGCTCTGAAG	313
QY	258	TACACTCTGCGCCGCAAGTTGGCGCTAGCTTCCGCGGCACTTGGCTTTTTCGCGTCCGAGGCT	317
Db	314	TACACTCTGCGCCGCAAGTTGGCGCTAGCTTCCGCGGCACTTGGCTTTTTCGCGTCCGAGGCT	373

Oy	318	IGTGTGCAAAAGGGCTCTTTGAGCTGCGCGCTGGCCGCGCGCTGGCCGAGCCGGAG	377
Db	374	TGTGTGCAAAAGGGCTTTTGAAGCTGCGCTTGGCCGCGCTGGCCGAGCCGGAG	433
Oy	378	GTCCCGAAGGGGGCTGCAGACCTGGCCCTTCGCGGAACTGGCCCGACGCGCG	437
Db	434	GTCCCGAAGGGGGCTGCAGACCTGGCCCTTCGCGGAACTGGCCCGACGCGCG	493
Oy	438	CGCAACCTTTCTATTTCAGGCTTCGCGGCGCTTTAGTACTCAGAGCGGAGCGAGA	497
Db	494	CGCAACCTTTCTATTTCAGGCTTCGCGGCGCTTTAGTACTCAGAGCGGAGCGAGA	553
Oy	498	GTAACAGGGCTGCAGCGCTTCTCTACCTGGCGTAGAGCTGGAGCTGGGACCCGACGCG	557
Db	554	GTAACAGGGCTGCAGCGCTTCTCTACCTGGCGTAGAGCTGGAGCTGGGAGCCGACGCGCG	613
Oy	558	GCGACAGCGCGGAGGGGAGCGCTGGAGAAAGCGAGCGGCGACGCGCGGGAGCCGAGATG	617
Db	614	GCGACAGCGCGGAGGGGAGCGCTGGAGAAAGCGAGCGGCGACGCGCGGGAGCCGAGATG	673
Oy	618	CAGCGGCGCGAAGCGGCGCGAGGTTTCGCGAGGGGAGCGGTGCGCCAGAGGTGAGAG	677
Db	674	CAGCGGCGCGAAGCGGCGCGAGGTTTCGCGAGGGGAGCGGTGCGCCAGAGGTGAGAG	733
Oy	678	CGCGCGCTCTGTCACTTGAAGCACTGTGGCGCTGCTCTCTCCGCTGGCCGAGACT	737
Db	734	CGCGCGCTCTGTCACTTGAAGCACTGTGGCGCTGCTCTCTCCGCTGGCCGAGAGT	793
Oy	738	TTCTGTGGCTCTGTTTGGGGCTGGCGCAAGGCGCGCTGGCGAATGAGCTTTGGCCACG	797
Db	794	TTCTGTGGCTCTGTTTGGGGCTGGCGCAAGGCGCGCTGGCGAATGAGCTTTGGCCACG	853
Oy	798	CCCTGCGCGGGGCGCCCTGTGTGACTGCTCCGAGCTGGCGCGCGCGCTGTGTC	857
Db	854	CCCTGCGCGGGGCGCCCTGTGTGACTGCTCCGAGCTGGCGCGCGCGCTGTGTC	913
Oy	858	TGGCGCCAGATTTTTGGAGTCCCTGGAGCCGGAACTTCCCGCTTGAGAGCCATGGGCG	917
Db	914	TGGCGCCAGATTTTTGGAGTCCCTGGAGCCGGAACTTCCCGCTTGAGAGCCATGGGCG	973
Oy	918	TCCACCTTGGGCTGCAGGCGCAGGAAACCACTCGCTGGAAATTGCAATTGTGGCTG	977
Db	974	TCCACCTTGGGCTGCAGGCGCAGGAAACCACTCGCTGGAAATTGCAATTGTGGCTG	1033
Oy	978	AAGTGTCCGCTGAAGTGAATGGGCGCATGTCAGAGATACCTCTTCCCGCCAGACATPA	1037
Db	1034	AAGTGTCCGCTGAAGTGAATGGGCGCATGTCAGAGATACCTCTTCCCGCCAGACATPA	1093
Oy	1038	CAGACAGCGGCTGTACATCTTCACTCTGGGACACAGGCGCTCCCGAAGGCTGTCGA	1097
Db	1094	CAGACAGCGGCTGTACATCTTCACTCTGGGACACAGGCGCTCCCGAAGGCTGTCGA	1153
Oy	1098	TCAGTCATCTGAAGTCTGTGCATGCCAGGCGTTCTATCAGCTGTGTGTCCACAGG	1157
Db	1154	TCAGTCATCTGAAGTCTGTGCATGCCAGGCGTTCTATCAGCTGTGTGTGTCCACAGG	1213
Oy	1158	AAAGTGTATCTACCTGCGCCCTCCCACTTACACATATCGGCTTCCGCTGGGCAATCG	1217
Db	1214	AAAGTGTATCTACCTGCGCCCTCCCACTTACACATATCGGCTTCCGCTGGGCAATCG	1273
Oy	1218	TGGGCTGATGGGCAATGGGGCCACAGTGTGTCTAAATCCAAAGTTTCGAGTGTGACT	1277
Db	1274	TGGGCTGATGGGCAATGGGGCCACAGTGTGTCTAAATCCAAAGTTTCGAGTGTGACT	1333
Oy	1278	TCTGGGAAGATTGGCAAGCAAGGTTGACGTTTCAAGTCAATTTGGGAGCTGTGCC	1337
Db	1334	TCTGGGAAGATTGGCAAGCAAGGTTGACGTTTCAAGTCAATTTGGGAGCTGTGCC	1393
Oy	1338	GATACCTTGTCAACACAGCCCCCGAGCAAGGCAAGATGAGCATAGGTTCGGGCTGGAG	1397
Db	1394	GATACCTTGTCAACACAGCCCCCGAGCAAGGCAAGATGAGCATAGGTTCGGGCTGGAG	1453
Oy	1398	TGGGCAAGCGGCTCGCCCAATACCTGGGAGCGTTTGTGTGCGGCGCTTCGGGCCCTTGC	1457

Db	1454	TGGCAGAGGGCGTGGCCAGATACCTGTGGAGCGTTTGTGGCGCGCTCGGGCCCTGCG	1513
QY	1458	AGGTGCTGGAGACATATGACTGACAGAGGCGAACGTGGCCACCATTAACATACACAGAC	1517
Db	1514	AGGTGCTGGAGACATATGACTGACAGAGGCGAACGTGGCCACCATTAACATACAGAGAC	1573
QY	1518	AGCGGGGCGCGTGGGCGGTGCTTCCTGGCTTTACAGACATATCTCCCTTCCTCTGA	1577
Db	1514	AGCGGGGCGCGTGGGCGGTGCTTCCTGGCTTTACAGACATATCTCCCTTCCTCTGA	1633
QY	1578	TTCCGTATGATGTACCAACAAGAGAGCCCATTCGGAGACCCCAAGGCGCACTGTATGCGCA	1637
Db	1634	TTCCGTATGATGTACCAACAAGAGAGCCCATTCGGAGACCCCAAGGCGCACTGTATGCGCA	1693
QY	1638	CATCTCCAGGTGAGCCAGGCGCTGCTGGGGCCCGGTAAGCCACAGTCCCATCTCTCGG	1697
Db	1634	CATCTCCAGGTGAGCCAGGCGCTGCTGGGGCCCGGTAAGCCACAGTCCCATCTCTCGG	1753
QY	1698	GCTATGCTGGCGGGCCAGAGCTGGCCCAAGGGGAAGTTGCTAAAGATGTCTTCCGGCTG	1757
Db	1754	GCTATGCTGGCGGGCCAGAGCTGGCCCAAGGGGAAGTTGCTAAAGATGTCTTCCGGCTG	1813
QY	1758	GGGATGTTTTTTTAAACACTGGGGGACCTGCTGGGTCTGGATGACCAAGTTTCTCGGCT	1817
Db	1814	GGGATGTTTTTTTAAACACTGGGGGACCTGCTGGGTCTGGATGACCAAGTTTCTCGGCT	1873
QY	1818	TCCATGATCTGTACTGAGACACCTTACAGGTGGAAGGGGGAGATGTGGCCCAACCGAGG	1877
Db	1874	TCCATGATCTGTACTGAGACACCTTACAGGTGGAAGGGGGAGATGTGGCCCAACCGAGG	1933
QY	1878	TGGCAGAGGCTTTGGAGGCCCTAGATTTTCTTACAGAGGTGAAGTCTATGAGTCACTG	1937
Db	1934	TGGCAGAGGCTTTGGAGGCCCTAGATTTTCTTACAGAGGTGAAGTCTATGAGTCACTG	1993
QY	1938	TGCCAGGGCATGAAAGGCGAGGCTGGAATGAGCCCTAGTCTCTCGTCCGCCCAACGCTT	1997
Db	1934	TGCCAGGGCATGAAAGGCGAGGCTGGAATGAGCCCTAGTCTCTCGTCCGCCCAACGCTT	2053
QY	1998	TGAGCTTATGACAGCTCTTACCCACCGTGTCTGAGAACTTGGCCACCTTATGGCCGGCCCT	2057
Db	2054	TGAGCTTATGACAGCTCTTACCCACCGTGTCTGAGAACTTGGCCACCTTATGGCCGGCCCT	2113
QY	2058	GATTCCTAGGCTCCAGGAGTCTTTGGCCACCAAGAGACCTTAAACAGCAGAAAGTTC	2117
Db	2114	GATTCCTAGGCTCCAGGAGTCTTTGGCCACCAAGAGACCTTAAACAGCAGAAAGTTC	2173
QY	2118	GGATGGCAATGAGGGCTTCGACCCCAACCTGTCTGACCCACTGTAGCTTCTGAGCC	2177
Db	2174	GGATGGCAATGAGGGCTTCGACCCCAACCTGTCTGACCCACTGTAGCTTCTGAGCC	2233
QY	2178	AGGCTGTAGGTGCTTACTGTCCTCTCAAACTTGCCCGGTATAGGGCCCTCTGGCAGGAA	2237
Db	2234	AGGCTGTAGGTGCTTACTGTCCTCTCAAACTTGCCCGGTATAGGGCCCTCTGGCAGGAA	2293
QY	2238	ACCTTCGATCTGAGAACTTCCACACTGAGAGCACTGAGAGAGAACTCTGTGGGTTGG	2297
Db	2234	ACCTTCGATCTGAGAACTTCCACACTGAGAGCACTGAGAGAGAACTCTGTGGGTTGG	2353
QY	2298	GGGCGCTTGCAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACAGAACTGCGGTCACT	2357
Db	2354	GGGCGCTTGCAGGTGTACTGGGCTGTCAAGGATCTTTTCTATACAGAACTGCGGTCACT	2413
QY	2358	ATTTGTATATATGTGCTGAGCTGATTCAGAGTCTCTGAGACTA 2404	
Db	2414	ATTTGTATATATGTGCTGAGCTGATTCAGAGTCTCTGAGACTA 2460	
RESULT 6			
AACT8481			
ID AACT8481 standard; cDNA, 2574 BP.			
XX AACT8481;			
AC AACT8481;			

XX 08-FEB-2001 (first entry)
 XX Human PRO703 (UNQ367) nucleotide sequence SEQ ID NO:101.
 XX Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
 KW expressed sequence tag; detection; cancer; ss.
 OS Homo sapiens.
 XX MO200053756-A2.
 XX 14-SEP-2000.
 XX 18-FEB-2000; 2000MO-US004341.
 XX 08-MAR-1999; 99MO-US005028.
 XX 12-MAR-1999; 99US-0123957P.
 XX 29-MAR-1999; 99US-0126773P.
 XX 21-APR-1999; 99US-0130232P.
 XX 28-APR-1999; 99US-0131445P.
 XX 14-MAY-1999; 99US-0134287P.
 XX 23-JUN-1999; 99US-0141037P.
 XX 26-JUL-1999; 99US-0145698P.
 XX 29-OCT-1999; 99US-0162506P.
 XX 30-NOV-1999; 99MO-US028313.
 XX 02-DEC-1999; 99MO-US028551.
 XX 16-DEC-1999; 99MO-US030095.
 XX 30-DEC-1999; 99MO-US031243.
 XX 05-JAN-2000; 2000MO-US000219.
 XX 06-JAN-2000; 2000MO-US000277.
 XX 06-JAN-2000; 2000MO-US000376.
 XX (GETH) GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Botstein D, Desnovers L, Eaton DL;
 PI Ferreira N, Filvarov E, Fong S, Gao W, Gerber H, Gettisen ME;
 PI Goddard AJ, Godowski PJ, Grimaldi CU, Gurney AL, Hillan KJ;
 PI Kijavini IU, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;
 PI Stewart TA, Tumas D, Williams PM, Wood WI;
 DR WPI: 2000-611443/58.
 DR P-PDB; AAB44255.
 XX Novel PRO polypeptides and polynucleotides used in detection methods, to
 PT target bioactive molecules to specific cells, and to modulate cellular
 PT activities.
 XX Claim 2; Fig 38; 636pp; English.
 XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed sequence
 CC tag) sequences which encode secreted or transmembrane PRO polypeptides.
 CC The PRO polynucleotides and polypeptides have cytostatic activity. The
 CC polynucleotides and polypeptides can be used for detecting the presence
 CC of PRO polypeptides in samples, for linking bioactive molecules to cells
 CC and for modulating biological activities of cells, using the polypeptides
 CC for specific targeting. The polypeptide targeting can be used to kill the
 CC target cells, e.g. for the treatment of cancers. The polypeptide pairs
 CC provide specific targeting of bioactive molecules to cells. AAC78600 to
 CC AAC78987 represent PCR primers and probes used in the isolation of the
 CC PRO polynucleotide sequences
 XX Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;
 SQ
 Query Match 99.3%; Score 2387; DB 3; Length 2574;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 CGCGCGCTCCCTGGAAGAGAAAGTCTCAGCTAGAAAGACCGCCCTAGGTTTCGGAAG 137
 DB 134 CGCGCGCTCCCTGGAAGAGAAAGTCTCAGCTAGAAAGACCGCCCTAGGTTTCGGAAG 193
 QY 138 GAGGATCAGGAGATGTTTTCAGAGCGGCTGAAACCAAGACGCTGCGATAGAGAACCGGCT 197
 DB 194 GAGGATCAGGAGATGTTTTCAGAGCGGCTGAAACCAAGACGCTGCGATAGAGAACCGGCT 253
 QY 198 CCATGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
 DB 254 CCATGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 313
 QY 258 TAACCTCTGCGCGCAGTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 317
 DB 314 TAACCTCTGCGCGCAGTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 373
 QY 318 TGTGCTGCAAAAGGCTCTTTCAGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 377
 DB 374 TGTGCTGCAAAAGGCTCTTTCAGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 433
 QY 378 GTCCCGAGGGGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
 DB 434 GTCCCGAGGGGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
 QY 438 CGACACCTTTCTCATTCACAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 497
 DB 494 CGACACCTTTCTCATTCACAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
 QY 498 GTAAACGGGCTGACAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
 DB 554 GTAAACGGGCTGACAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
 QY 558 GCGACAGCGCGGAGGAGAGCGCTGAGAAAGCGAGCGGCGAGCGCGGAGCGGAGATG 617
 DB 614 GCGACAGCGCGGAGGAGAGCGCTGAGAAAGCGAGCGGCGAGCGGAGCGGAGATG 673
 QY 618 CACGCGCGGAGAGCGGCGGAGATTTGCGGAGGAGGAGCGGCTGCGGAGGAGGAGGAGG 677
 DB 674 CACGCGCGGAGAGCGGCGGAGATTTGCGGAGGAGGAGCGGCTGCGGAGGAGGAGGAGG 733
 QY 678 CCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
 DB 734 CCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
 QY 738 TTCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 797
 DB 794 TTCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 853
 QY 798 CCTGCGCGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857
 DB 854 CCTGCGCGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
 QY 858 TGGCGCGAAGTTCGAGTCCCTGAGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 917
 DB 914 TGGCGCGAAGTTCGAGTCCCTGAGAGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973
 QY 918 TCCACCTGAGGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 977
 DB 974 TCCACCTGAGGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1033
 QY 978 AAGTGTCCGCTGAGAGTGGGCGGAGTCCAGAGTACCTCTTCCGCCAGAGCATTA 1037
 DB 1034 AAGTGTCCGCTGAGAGTGGGCGGAGTCCAGAGTACCTCTTCCGCCAGAGCATTA 1093
 QY 1038 CAGACAGTGGCTGTATCATCTTCACTTCTGAGACAGACAGGCGCTCCGCAAGGCTGTCCGA 1097
 DB 1094 CAGACAGTGGCTGTATCATCTTCACTTCTGAGACAGACAGGCGCTCCGCAAGGCTGTCCGA 1153
 QY 1098 TCAGTATCTGAAGATCTGCAATGCAATGCAAGGCTTCTATCAGCTGTGTGTGTCCACAGG 1157
 DB 1154 TCAGTATCTGAAGATCTGCAATGCAATGCAAGGCTTCTATCAGCTGTGTGTGTCCACAGG 1213

QY	1158	AAAGTGTGATCTACCTGACCCCTCCCACTCAACCAAGTCAGGTTTCCCTGCGGACATCG	1217
Db	1214	AAAGTGTGATCTACCTGACCCCTCCCACTCTCAACCAAGTCAGGTTTCCCTGCGGACATCG	1273
QY	1218	TGGGCTGCGATGGGCATTGGGAGCCACAGTGTGCTGAATCCAAAGTTCTCGGCTGCTCAGT	1277
Db	1274	TGGGCTGCGATGGGCATTGGGAGCCACAGTGTGCTGAATCCAAAGTTCTCGGCTGCTCAGT	1333
QY	1278	TTGCGGAAGAATTGCGACAGACAAGGGTGAACGGTTCCAGTAACTTGGGGAAGCTGGCC	1337
Db	1334	TTGCGGAAGAATTGCGACAGACAAGGGTGAACGGTTTCAGTAACTTGGGGAAGCTTGGCC	1393
QY	1338	GATACCTTTGTCACACACACCCCGAGCAAGCAGACGTGGCCATTAAGTTCGGCTGGCAG	1397
Db	1394	GATACCTTTGTCACACACGCCCCGAGCAAGCAGAACGTGGCCATTAAGTTCGGCTGGCAG	1453
QY	1398	TGGGCAAGCGGCGCTGCGGCCAGATACCTTGGGAGGCTTTTGTGCGGGCGCTTGGGCGCCCTGC	1457
Db	1454	TGGGCAACCGGGCTGCGGCCAGATACCTTGGGAGGCTTTTGTGCGGGCGCTTGGGCGCCCTGC	1513
QY	1458	AGGTGCTGGAACAATATGACTGACAGAGGCAACGTGGCCACATCAACTACACAGAAC	1517
Db	1514	AGGTGCTGGAACAATATGACTGACAGAGGCAACGTGGCCACATCAACTACACAGAAC	1573
QY	1518	AGCGGGGCGCTGTGAGGCGCTGCTTCTGGCTTTACAAACATATCTTCCCTTCTCTTGA	1577
Db	1574	AGCGGGGCGCTGTGAGGCGCTGCTTCTGGCTTTACAAACATATCTTCCCTTCTCTTGA	1633
QY	1578	TTGGCTTATGATGACACACACAGAGAGCCAAATGGGGAGCCCCAGGGGCACTGTATGGCA	1637
Db	1634	TTGGCTTATGATGACACACAGAGAGCCAAATGGGGAGCCCCAGGGGCACTGTATGGCA	1693
QY	1638	CATCTCCAGTGTGAGCCAGGCGCTGTGTGTGGCCCGGTAAAGCCAGACAGTCCCATTTCTCG	1697
Db	1694	CATCTCCAGTGTGAGCCAGGCGCTGTGTGTGGCCCGGTAAAGCCAGACAGTCCCATTTCTCG	1753
QY	1698	GCTATGCTGCGGGGCGAGAGCTGGCCAGGGGAAAGTTGCTTAAAGATGTTCTTCCGGGCTG	1757
Db	1754	GCTATGCTGCGGGGCGAGAGCTGGCCAGGGGAAAGTTGCTTAAAGATGTTCTTCCGGGCTG	1813
QY	1758	GGAAATGTTTTCTTCAACACTGAGGACCTGCTGATGCTGAGTGAACCAAGTTTCTCCGCT	1817
Db	1814	GGAAATGTTTTCTTCAACACTGAGGACCTGCTGATGCTGAGTGAACCAAGTTTCTCCGCT	1873
QY	1818	TCCATGATCGTACTGAGACACCTTCAGTGGAGAGGGGAGAAATGTGGCCACAAACGAGG	1877
Db	1874	TCCATGATCGTACTGAGACACCTTCAGTGGAGAGGGGAGAAATGTGGCCACAAACGAGG	1933
QY	1878	TGGCAGAGTCTTOSAGGCGCCCTAGATTTTCTTCAAGAGGTGAAGCTCATATGAGATCATG	1937
Db	1934	TGGCAGAGTCTTOSAGGCGCCCTAGATTTTCTTCAAGAGGTGAAGCTCATATGAGATCATG	1993
QY	1938	TGCCAGGCGCATGAAGGCAAGGCGTGGAAATGGCAAGCCCTATGCTGCGTCCCCCACGCTT	1997
Db	1994	TGCCAGGCGCATGAAGGCAAGGCGTGGAAATGGCAAGCCCTATGCTGCGTCCCCCACGCTT	2053
QY	1998	TGACACTTATGACGCTCTTACACCCACGTGTCTGAAGAACTTGGCACTTATATGCCGGGCCC	2057
Db	2054	TGACACTTATGACGCTCTTACACCCACGTGTCTGAAGAACTTGGCACTTATATGCCGGGCCC	2113
QY	2058	GATTCCTCAGGCTCAGAGAGTCTTGGGCAACACAGAGACCTTCAACAGCAGAAAGTTT	2117
Db	2114	GATTCCTCAGGCTCAGAGAGTCTTGGGCAACACAGAGACCTTCAACAGCAGAAAGTTT	2173
QY	2118	GGATGGCAATAGAGGCTTCCAGCCCGACACCTCTGTGAACCACTTATGCTTCTGAGAC	2177
Db	2174	GGATGGCAATAGAGGCTTCCAGCCCGACACCTCTGTGAACCACTTATGCTTCTGAGAC	2233
QY	2178	AGGCTGTAGGTCCTTACCTGCCCCCTTCAACAATGCCCGGATCAGGCCCTCTCTGAGAGAA	2237
Db	2234	AGGCTGTAGGTCCTTACCTGCCCCCTTCAACAATGCCCGGATCAGGCCCTCTCTGAGAGAA	2293
QY	2238	ACCTTCGAATCTGAGAACTTCCACACTGAGGCACTGTAGAGAGAACTCTGTGGGATGG	2297

Db	2294	ACCTTGGAATCTGAGAACTTCACACCTGAGGACCGTAGAGAGGAACCTGTGGGGGTGG	2353
Qy	2298	GGGGCCGTGAGGAGTACTGAGGCTGTGAGGAGATCTTTCTATACCAAGATGCGGTCACT	2357
Db	2354	GGGGCCGTGAGGAGTACTGAGGCTGTGAGGAGATCTTTCTATACCAAGATGCGGTCACT	2413
Qy	2358	ATTTTGTAATAAATGTGGCTGAGGCTGATCCAGCTGTCTGTGACCTA	2404
Db	2414	ATTTTGTAATAAATGTGGCTGAGGCTGATCCAGCTGTCTGTGACCTA	2460
RESULT 7			
AAcS8239			
ID	AAcS8239	standard; cDNA; 2574 BP.	
XX			
AC	AAcS8239;		
XX			
DT	25-JAN-2001	(first entry)	
XX			
DE	Human PRO703	nucleotide sequence SEQ ID NO:28.	
XX			
KW	Human; tumour; diagnosis; neoplastic disease; identification; cancer;		
KW	tumorigenesis; detection; neoplastic cell growth; proliferation;		
KW	cytoskeletal; anti-inflammatory; immunomodulatory; inflammatory disorder;		
KM	Immunological disorder; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200053754-A1.		
XX			
PD	14-SEP-2000.		
XX			
PF	06-JAN-2000; 2000WO-US900277.		
XX			
PR	08-MAR-1999; 99WO-US005028.		
PR	12-MAR-1999; 99US-0123957P.		
PR	29-MAR-1999; 99US-0126773P.		
PR	21-APR-1999; 99US-0130232P.		
PR	28-APR-1999; 99US-0131445P.		
PR	05-OCT-1999; 99WO-US023089.		
PR	30-NOV-1999; 99WO-US028551.		
PR	02-DEC-1999; 99WO-US028551.		
PR	02-DEC-1999; 99WO-US028564.		
PR	30-DEC-1999; 99WO-US031243.		
PR	30-DEC-1999; 99WO-US031274.		
XX			
PA	(GETH) GENENTECH INC.		
XX			
PI	Baker KP, Desauwage FJ, Goddard A, Gurney AL, Klein RD, Roy MA;		
PI	Wood WJ;		
XX			
DR	WPI; 2000-572269/53.		
XX			
P	P-PEDB; AAB24054.		
XX			
PT	New isolated antibody for use in compositions and methods for the		
PT	diagnosis and treatment of neoplastic cell growth and proliferation in		
PT	mammals, including humans, and in monitoring tumor treatment.		
XX			
XX	Claim 50; Fig 28; 195pp; English.		
XX			
XX	The present invention describes an isolated antibody (Ab) that binds to		
CC	one of the human proteins (P) designated PRO213, PRO1330, PRO1449,		
CC	PRO327, PRO324, PRO351, PRO615, PRO531, PRO538, PRO3664, PRO618,		
CC	PRO772, PRO703, PRO792 or PRO474. The Ab can be used in compositions and		
CC	methods for the diagnosis and treatment of neoplastic cell growth and		
CC	proliferation in mammals, including humans. Genes and polypeptides		
CC	encoded by them, that are amplified in the genome of a tumour cell, can		
CC	be identified and are useful targets for the treatment and prevention of		
CC	certain cancers and may be used to monitor tumour treatment. Compounds		
CC	that inhibit the expression or activity of the identified polypeptides		
CC	can be identified and used as antagonists. Benign or malignant tumours,		
CC	inflammatory disorders and immunological disorders can be treated.		

DB 974 TCCACCTGTGGGCGTCGACGGCCACAGAAACCACCTGCTGGATTAGAGCATTTGTGGGTG 1033

QY 978 AAGTGTCCGCTGAAGTGGATGGCCAGTGCAGGATACCTTCTTCCCCCGACGATAA 1037

DB 1034 AAGGTCCGCTGAAGTGGATGGCCAGTGCAGGATACCTTCTTCCCCCGACGATAA 1093

Db 1094 CAGCACGTCCTGTACATCTTCACCTGTGGCACACCGGGCTTCCCAAGGCTGTGGGA 1153

Db	1154	TGACATCATCTGAAATCTCTGCAATGCCAGGCTCTATCAGCTGTGTGTCCACAGG	1213
Oy	1158	AAGATGTATCTACCTTGCCCTCCACTCTACACATATCTCGGTTCCCTGTGSGCATCG	1217
Db	1214	AAGATGTGATCTACCTGCGCCCTCCACTCTACACATATCTCGGTTCCCTGTGSGCATCG	1273
Oy	1218	TGGCTCTCATGGGCAATGGGGGCCACAGTGGGTGTGAATCCAAATTTCCGCTGTGTAGT	1277
Db	1274	TGGCTCTCATGGGCAATGGGGGCCACAGTGTGTGAATCCAAATTTCCGCTGTGTAGT	1333
Oy	1278	TTGTGGAAAGATTGCGACAGACAGGGGTGAACGTTCTCAAGTACATTGGGAGCTGTGCC	1337
Db	1334	TTGTGGAAAGATTGCGACAGACAGGGGTGAACGTTCTCAAGTACATTGGGAGCTGTGCC	1393
Oy	1338	GATACCTTGTCAACACAGCCCCGAGCAAGGCAAGACGTGGCCATTAAGGTCCGCTGGCAG	1397
Db	1394	GATACCTTGTCAACACAGCCCCGAGCAAGGCAAGACGTGGCCATTAAGGTCCGCTGGCAG	1453
Oy	1398	TGGGCAAGGGGCTGCGCCAGATPACTGGGAGCGTTTGTGCGCGCTTGGGCGCTCTGC	1457
Db	1454	TGGGCAAGGGGCTGCGCCAGATPACTGGGAGCGTTTGTGCGCGCTTGGGCGCTCTGC	1513
Oy	1458	AGGTGCTGAGACATATGACTGACAGAGGGCAACGTGGCCACATCACTACAGAGAC	1517
Db	1514	AGGTGCTGAGACATATGACTGACAGAGGGCAACGTGGCCACATCACTACAGAGAC	1573
Oy	1518	AGCGGGGCGCTGTGGGGCGTGTCTTCTGCGCTTTACAAGCATPCTTCCCTTCTCTTGA	1577
Db	1574	AGCGGGGCGCTGTGGGGCGTGTCTTCTGCGCTTTACAAGCATPCTTCCCTTCTCTTGA	1633
Oy	1578	TTGCGCTATGATGTACACACAGAGAGGCAATTGGGGAGCCCCAGGGGCACTGATGSGCA	1637
Db	1634	TTGCGCTATGATGTACACACAGAGAGGCAATTGGGGAGCCCCAGGGGCACTGATGSGCA	1693
Oy	1638	CATCTCCAGGTGAGCCAGGCGTGTGTGTGGCCCGGTAAAGCCAGCATGTCCCATTTCTGTG	1697
Db	1694	CATCTCCAGGTGAGCCAGGCGTGTGTGTGGCCCGGTAAAGCCAGCATGTCCCATTTCTGTG	1753
Oy	1698	GCTATGCTGGCGGAGCCAGAGCTGGCCAGGGGAAAGTGTCTAAAGATATGTTCCGGGCTG	1757
Db	1754	GCTATGCTGGCGGAGCCAGAGCTGGCCAGGGGAAAGTGTCTAAAGATATGTTCCGGGCTG	1813
Oy	1758	GGAATGTTTTCTTCAACACTGGGGGACCTGTGTGTCTGCGATGACCAAGTTTTCTCCGCT	1817
Db	1814	GGAATGTTTTCTTCAACACTGGGGGACCTGTGTGTCTGCGATGACCAAGTTTTCTCCGCT	1873
Oy	1818	TTCATGATCGTACTGAGACACCTTCAGGTGGAAGGGGAGATGTGGCCACACAGAG	1877
Db	1874	TTCATGATCGTACTGAGACACCTTCAGGTGGAAGGGGAGATGTGGCCACACAGAG	1933
Oy	1878	TGGCAAGAGTCTTCGAGGCGCTTAGATTTTCTTCAAGAGGTGAACGTCTATGAGTCACTG	1937
Db	1934	TGGCAAGAGTCTTCGAGGCGCTTAGATTTTCTTCAAGAGGTGAACGTCTATGAGTCACTG	1993
Oy	1938	TGGCAGGGCATGAAGGAGGGGCTGGAATGGACAGCTAGTCTGCGTCCGCCACAGCTT	1997
Db	1994	TGGCAGGGCATGAAGGAGGGGCTGGAATGGACAGCTAGTCTGCGTCCGCCACAGCTT	2053
Oy	1998	TGGACTTATGACGCTCTACACCCACGTGTGTGAAGACTTGGCACTTATGCCCGGCC	2057
Db	2054	TGGACTTATGACGCTCTACACCCACGTGTGTGAAGACTTGGCACTTATGCCCGGCC	2113

```
OY 2058 GATTCTCAGGCTTCAGAGCTTTGGCCACACAGAGACCTTCAACAGCAAGTTTC 2117
Db 2114 GATTCTCAGGCTTCAGAGCTTTGGCCACACAGAGACCTTCAACAGCAAGTTTC 2173
OY 2118 GGATGCAAAATGAGGGCTTCGACCCAGACACCTGTCTGCAACCTGTACGTTCTGACC 2177
Db 2174 GGATGCAAAATGAGGGCTTCGACCCAGACACCTGTCTGCAACCTGTACGTTCTGACC 2233
OY 2178 AGGCTGAGGCTCTTACTGCGCCCTTCACAACTGCCGCTACAGGCCCTCTGCGAGGA 2237
Db 2234 AGGCTGAGGCTCTTACTGCGCCCTTCACAACTGCCGCTACAGGCCCTCTGCGAGGA 2293
OY 2238 ACCTTCGATCTGAGAACTTCCACACTGAGGACCTGAGAGAGAACTGTGTGGGTGG 2297
Db 2294 ACCTTCGATCTGAGAACTTCCACACTGAGGACCTGAGAGAGAACTGTGTGGGTGG 2353
OY 2298 GGGCCGTTCAGGCTGTACTGGGCTGTCAAGGATCTTTTCTATACCAAGACTGCGGTACT 2357
Db 2354 GGGCCGTTCAGGCTGTACTGGGCTGTCAAGGATCTTTTCTATACCAAGACTGCGGTACT 2413
OY 2358 ATTTGTATTAAGTGTGCTGAGACTGATCCAGCTGTCTGACCTA 2404
Db 2414 ATTTGTATTAAGTGTGCTGAGACTGATCCAGCTGTCTGACCTA 2460

RESULT 8
ACD42510
ID ACD42510 standard; cDNA, 2574 BP.
XX ACD42510;
AC ACD42510;
DT 09-SEP-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO703 cDNA.
XX
XX Human, secreted and transmembrane protein; PRO; virucide; gene therapy;
KM cell death; growth induction cascade; blood coagulation cascade;
KM viral infection; gene; ss.
XX
XX Homo sapiens.
OS
XX
PN US2003050239-A1.
XX
PD 13-MAR-2003.
XX
PF 15-OCT-2001; 2001US-00978191.
XX
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077793P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-0004022P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0078984P.
PR 26-MAR-1998; 98US-0078656P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 30-MAR-1998; 98US-0079820P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.
```

```
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080333P.
PR 01-APR-1998; 98US-0080334P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 09-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081203P.
PR 09-APR-1998; 98US-0081229P.
PR 15-APR-1998; 98US-0081817P.
PR 15-APR-1998; 98US-0081819P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0081952P.
PR 15-APR-1998; 98US-0081955P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 22-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083336P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084418P.
PR 06-MAY-1998; 98US-0084498P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 13-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085573P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085822P.
PR 15-MAY-1998; 98US-0085889P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086332P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086450P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087088P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0030863P.
PR 26-JUN-1998; 98US-0031010P.
PR 01-JUL-1998; 98US-0031359P.
PR 30-JUL-1998; 98US-0034651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-00168978.
```


PR 07-OCT-1998; 98WO-US021141.
 PR 02-NOV-1998; 98US-00184216.
 PR 06-NOV-1998; 98US-00187368.
 PR 20-NOV-1998; 98US-0109304P.
 PR 07-DEC-1998; 98WO-US024855.
 PR 22-DEC-1998; 98US-00202054.
 PR 22-DEC-1998; 98US-00218517.
 PR 23-DEC-1998; 98US-0113296P.
 PR 05-JAN-1999; 98US-0113621P.
 PR 05-MAR-1999; 99WO-US000106.
 PR 08-MAR-1999; 99US-00254465.
 PR 10-MAR-1999; 99WO-US005028.
 PR 10-MAR-1999; 99US-00265686.
 PR 10-MAR-1999; 99WO-US005190.
 PR 12-MAR-1999; 99US-00267213.
 PR 12-MAR-1999; 99US-0123957P.
 PR 29-MAR-1999; 99US-0126773P.
 PR 12-APR-1999; 99US-00284291.
 PR 21-APR-1999; 99US-0130232P.
 PR 26-APR-1999; 99US-0131022P.
 PR 28-APR-1999; 99US-0131445P.
 PR 14-MAY-1999; 99US-00311832.
 PR 14-MAY-1999; 99US-0134287P.
 PR 02-JUN-1999; 99WO-US010733.
 PR 16-JUN-1999; 99WO-US012252.
 PR 23-JUN-1999; 99US-0139557P.
 PR 07-JUL-1999; 99US-0142680P.
 PR 26-JUL-1999; 99US-0145698P.
 PR 28-JUL-1999; 99US-0146222P.
 PR 25-AUG-1999; 99US-00380137.
 PR 25-AUG-1999; 99US-00380138.
 PR 25-AUG-1999; 99US-00380142.
 PR 29-OCT-1999; 99US-0162506P.
 PR 30-NOV-1999; 99WO-US028313.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 10-MAR-2000; 2000WO-US006319.
 PR 21-MAR-2000; 2000WO-US007352.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014941.
 PR 30-MAY-2000; 2000WO-US015264.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 24-AUG-2000; 2000WO-US023328.
 PR 08-NOV-2000; 2000US-00709238.
 PR 27-NOV-2000; 2000US-00723749.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000US-00747259.
 PR 20-DEC-2000; 2000WO-US04956P.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 22-MAR-2001; 2001US-00816744.
 PR 22-MAR-2001; 2001US-00816920.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854280.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001WO-US017880.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.

PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 30-JUL-2001; 2001US-00918585.
 XX
 XX
 PA (GETH) GENENTECH INC.
 PI Ashkenazi AJ, Baker KP, Botstein D, Deenoyers L, Eaton DL,
 PI Ferrera N, Filvaroff E, Fong S, Gao W, Gerber H, Gerltzen ME;

Query Match 99.3%; Score 2387; DB 7; Length 2574;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 CTGCTCCGCGCGGTGAGTGTGGGCGCTGAGTGGAAATGGCGCTGTGCACGCA 77
 Db 74 CTGCTCCGCGCGGTGAGTGTGGGCGCTGAGTGGAAATGGCGCTGTGCACGCA 133
 QY 78 CGGCGCTCCCTGAAAGGAATCTGAGTGAAGAGCGCGCTTGTGGAGG 137
 Db 134 CGGCGCTCCCTGAAAGGAATCTGAGTGAAGAGCGCGCTTGTGGAGG 193
 QY 138 GAGATCAGGAGTGTGGAGCGGCTGGAACGAGCGGTGCGATAGAGAGCGGCT 197
 Db 194 GAGATCAGGAGTGTGGAGCGGCTGGAACGAGCGGTGCGATAGAGAGCGGCT 253
 QY 198 CCAATGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
 Db 254 CCAATGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 313
 QY 258 TACACCTGCGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 317
 Db 314 TACACCTGCGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 373
 QY 318 TGTGCTGCAAAAGGCTCTTGAAGTCCGCGCTGCGGCTGCGGCTGCGGCTGCGGCT 377
 Db 374 TGTGCTGCAAAAGGCTCTTGAAGTCCGCGCTGCGGCTGCGGCTGCGGCTGCGGCT 433
 QY 378 GTCCGAGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 437
 Db 434 GTCCGAGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 493
 QY 438 CGCACACTTTTCATTCACGCGCTGCGGCTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCT 497
 Db 494 CGCACACTTTTCATTCACGCGCTGCGGCTGCGGCTTGTGCTGCTGCTGCTGCTGCTGCT 553
 QY 498 GTTACAGGCTGCAAGCGGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
 Db 554 GTTACAGGCTGCAAGCGGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
 QY 558 GCGACAGCGGCGAGGAGCGCTGAGAGAGCGAGCGGCGAGCGGCGAGCGGCGAGATG 617
 Db 614 GCGACAGCGGCGAGGAGCGCTGAGAGAGCGAGCGGCGAGCGGCGAGCGGCGAGATG 673
 QY 618 CAGCGCGCGAAGCGCGCGGAGTTGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 677
 Db 674 CAGCGCGCGAAGCGCGCGGAGTTGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 733
 QY 678 CGCGCGCGCTTGTACCTGAGAGCACTGTGAGCGCTCTCTCCCGCTGCGGCGGCGGCT 737
 Db 734 CGCGCGCGCTTGTACCTGAGAGCACTGTGAGCGCTCTCTCCCGCTGCGGCGGCGGCT 793
 QY 738 TTCTGTGCTGTGTTGAGGCTGAGCGGCGGCGGCTGAGCGGCTTGTGCGGCGGCGGCT 797
 Db 794 TTCTGTGCTGTGTTGAGGCTGAGCGGCGGCGGCTTGTGCGGCGGCGGCTTGTGCGGCGG 853
 QY 798 CCTGCGCGCGGCGGCGGCTTGTGCGGCTTGTGCGGCTTGTGCGGCGGCGGCGGCGGCT 857
 Db 854 CCTGCGCGCGGCGGCGGCTTGTGCGGCTTGTGCGGCTTGTGCGGCGGCGGCGGCGGCT 913
 QY 858 TGGCGCGAGGTTTGTGAGTCCCTGAGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCT 917
 Db 914 TGGCGCGAGGTTTGTGAGTCCCTGAGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCT 973

QY	918	TTCCACCTGTTGGGCTGCAAGGCGCCACAGAAACCCACCCTGTGTGAAATTAAGGATTTTCTGGCTG	977
Db	974	TTCCACCTGTGGGCTGCAAGGCGCCACAGAAACCCACCCTGTGTGAAATTAAGGATTTTCTGGCTG	1033
QY	978	AAAGTTCGGCTGGAAGTGGATGGGCACTGGCCAGAGTATCCTCTCTTCCCTCCCAAGACATTA	1037
Db	1034	AAAGTTCGGCTGGAAGTGGATGGGCACTGGCCAGAGTATCCTCTCTTCCCTCCCAAGACATTA	1093
QY	1038	CAGACACATGTCCTTATCATCTTCACTCTGTGGACACAAGGAGCTTCCCAAGGCTGTGGAA	1097
Db	1094	CAGACACATGTCCTTATCATCTTCACTCTGTGGACACAAGGAGCTTCCCAAGGCTGTGGAA	1153
QY	1098	TCACTGCATCTGAAGATCCTGTGCATATGCCAGGGCTTCTATACAGCTGTGTGTGTCCACCAAG	1157
Db	1154	TCACTGCATCTGAAGATCCTGTGCATATGCCAGGGCTTCTATACAGCTGTGTGTGTCCACCAAG	1213
QY	1158	AAAGTGTATGTATCCTGGCCCTCCCACTCTACACACATGTCCGGTTCCTCTGCTGGGACATCG	1217
Db	1214	AAAGTGTATGTATCCTGGCCCTCCCACTCTACACACATGTCCGGTTCCTCTGCTGGGACATCG	1273
QY	1218	TGGCTGTGATGGGCAATTGGGSCCAACAGTGTGTCTGAATTCACAGTTCTCGGCTGTGTACT	1277
Db	1274	TGGCTGTGATGGGCAATTGGGSCCAACAGTGTGTCTGAATTCACAGTTCTCGGCTGTGTACT	1333
QY	1278	TCTGGGAAAGATTGGCCAGACAGCAAGGGGTGATCGGTTCACATTAATTGGGAGCTGTGCC	1337
Db	1334	TCTGGGAAAGATTGGCCAGACAGCAAGGGGTGATCGGTTCACATTAATTGGGAGCTGTGCC	1393
QY	1338	GATACCTTTGTCAACCAAGCCCCGAGCAAGGCAGAAAGTGGCCATPAAGTCCGGCTGGCAG	1397
Db	1394	GATACCTTTGTCAACCAAGCCCCGAGCAAGGCAGAAAGTGGCCATPAAGTCCGGCTGGCAG	1453
QY	1398	TGGGAGAGGGGCTGTGGCCACAGATACCTGGAGCGTTTGTCCGGCGCTCGGCGCCCTGCG	1457
Db	1454	TGGGAGAGGGGCTGTGGCCACAGATACCTGGAGCGTTTGTGTGGCGCGTTTGTGGCGCCCTGCG	1513
QY	1458	AGGTGCTGGAGACATATGAGTCTGACAGAGGCGCAAGTGGCCACCATCACTACACAGAGC	1517
Db	1514	AGGTGCTGGAGACATATGAGTCTGACAGAGGCGCAAGTGGCCACCATCACTACACAGAGC	1573
QY	1518	AGCGGGGGGCGTGTGGGCGGTGCTTCTGAGCTTTTCAACATATCTTCCCTTCTCCTTGA	1577
Db	1574	AGCGGGGGGCGTGTGGGCGGTGCTTCTGAGCTTTTCAACATATCTTCCCTTCTCCTTGA	1633
QY	1578	TTTCGTATGATGTCAACAAGAGAGCCAAATTGGGAAATCCCAAGGGCACTGTATGTCGA	1637
Db	1634	TTTCGTATGATGTCAACAAGAGAGCCAAATTGGGAAATCCCAAGGGCACTGTATGTCGA	1693
QY	1638	CATTCGCCAGGTAGGCAAGGAGGCTGTGTGTGGCCCGGTAAAGCAGACAGTCCCATTTCTGTG	1697
Db	1694	CATTCGCCAGGTAGGCAAGGAGGCTGTGTGTGGCCCGGTAAAGCAGACAGTCCCATTTCTGTG	1753
QY	1698	GCTATGTCTGGCGGGCCAGAGCTGTGCCCAAGGGGAATGTCTTAAGGATGTCTTCGGGCTGT	1757
Db	1754	GCTATGTCTGGCGGGCCAGAGCTGTGCCCAAGGGGAATGTCTTAAGGATGTCTTCGGGCTGT	1813
QY	1758	GGGATGTGTTTCTTCAACAACGTGGGGAACCTGTGTGTGTGCATGACCAAGATTTTCTCCGCT	1817
Db	1814	GGGATGTGTTTCTTCAACAACGTGGGGAACCTGTGTGTGTGCATGACCAAGATTTTCTCCGCT	1873
QY	1818	TCATATGATGTGACTGGAGACACCTTCAGGTGGAAAGGGGGAATGTGGCCACACACCGAG	1877
Db	1874	TCATATGATGTGACTGGAGACACCTTCAGGTGGAAAGGGGGAATGTGGCCACACACCGAG	1933
QY	1878	TGGCAGAGGCTCTTCCAGGCGCTTAGATTTTCTTCAGAGGTGGAACGTCTATGTAGATCACTG	1937
Db	1934	TGGCAGAGGCTCTTCCAGGCGCTTAGATTTTCTTCAGAGGTGGAACGTCTATGTAGATCACTG	1993
QY	1938	TGGCAGAGGCAATGAAGCCAGGCGCTGGAATGCAAGCCCTAATTTGTGGTCCCCCCCCACGCTT	1997
Db	1994	TGGCAGAGGCAATGAAGCCAGGCGCTGGAATGCAAGCCCTAATTTGTGGTCCCCCCCCACGCTT	2053

QY	1998	TGACACCTTAATGACAGCTCTTAACCCACAGTGTCTAGAACTTGGCACCTTAATGACCGGCC	2057
Db	2054	TGACCTTAATGACAGCTCTTAACCCACGATGTCTGAACTTGGCACCTTAATGACCGGCC	2113
QY	2058	GATTCCTCAGAGCTCCAGAGATCTTTGGCCACCAACAGACCTTCAACAGCAGAAAGTTC	2117
Db	2114	GATTCCTCAGAGCTCCAGAGATCTTTGGCCACCAACAGAGACCTTCAACAGCAGAAAGTTC	2173
QY	2118	GGATGCGAAATGAGGGCTTTGACCCCGACACCTCTGTGACCCCACTGATGAGTCTGAGAC	2177
Db	2174	GGATGCGAAATGAGGGCTTTGACCCCGACACCTCTGTGACCCCACTGATGAGTCTGAGAC	2233
QY	2178	AGGCTGTAGAGTGCCTACCTGCCCTCCACAACTGCCCGATCAGGCCCTCTCTGACAGAA	2237
Db	2234	AGGCTGTAGAGTGCCTACCTGCCCTCCACAACTGCCCGATCAGGCCCTCTCTGACAGAA	2293
QY	2238	ACCTTCGAAATCTGAGAACTTCCACACTTGAGGCACTTGAGAGAGAACTCTGTGGGGTGG	2297
Db	2294	ACCTTCGAAATCTGAGAACTTCCACACTTGAGGCACTTGAGAGAGAACTCTGTGGGGTGG	2353
QY	2298	GGGCGCGTGTGACGATGATCTGGGCGTGTACAGGGATCTTTTCTATACAGAACTGCGATCACT	2357
Db	2354	GGGCGCGTGTGACGATGATCTGGGCGTGTACAGGGATCTTTTCTATACAGAACTGCGATCACT	2413
QY	2358	ATTTTGTATATATATGTGGCTGAGACTGATCCAGCTGTCTGACCTA	2404
Db	2414	ATTTTGTATATATATGTGGCTGAGACTGATCCAGCTGTCTGACCTA	2460
RESULT 9			
ID	ACA63545		
AC	ACA63545 standard; cDNA; 2574 BP.		
XX	ACA63545;		
XX	16-JUN-2003 (first entry)		
DE	Novel human secreted and transmembrane protein PRO703 cDNA.		
KM	Human; secreted and transmembrane protein; PRO; antiinflammatory;		
KM	antiartherosclerotic; cardiant; anti-infectivity; anti-HIV; cytostatic;		
KM	antidiabetic; gene therapy; inflammatory disease; organ failure;		
KM	atherosclerosis; cardiac injury; infertility; birth defect;		
KM	premature aging; AIDS; cancer; diabetic complication; chromosome mapping;		
KM	gene mapping; pharmaceuticals; diagnostic; biosensor; bioreactor;		
KM	tissue typing; gene; ss.		
OS	Home sapiens.		
XX	US2002192706-A1.		
XX	19-DEC-2002.		
PD	24-OCT-2001; 2001US-00999832.		
PF			
XX	17-OCT-1997; 97US-0062250P.		
PR	03-NOV-1997; 97US-0064249P.		
PR	21-NOV-1997; 97US-0065311P.		
PR	11-NOV-1997; 97US-0066364P.		
PR	10-MAR-1998; 98US-0077450P.		
PR	11-MAR-1998; 98US-0077632P.		
PR	11-MAR-1998; 98US-0077641P.		
PR	11-MAR-1998; 98US-0077649P.		
PR	12-MAR-1998; 98US-0077791P.		
PR	13-MAR-1998; 98US-0078004P.		
PR	17-MAR-1998; 98US-00040220.		
PR	20-MAR-1998; 98US-0078886P.		
PR	20-MAR-1998; 98US-0078910P.		
PR	20-MAR-1998; 98US-0078936P.		
PR	20-MAR-1998; 98US-0078939P.		
PR	25-MAR-1998; 98US-0079294P.		
PR	26-MAR-1998; 98US-0079656P.		
PR	27-MAR-1998; 98US-0079663P.		

	PI	Ashkenazi I ¹ , Baker KP, Boesein D, Desnoyers L, Eaton DL;
	PI	Fernandez N, Filyarov E, Pong S, Gao W, Garber H, Gerritsen MB;
	P1	Goddard A, Goodwest PJ, Grimaldi JC, Gurrey AL, Hillan KJ,
	P1	Kijavittu J, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DJ;
	P1	Stewart TA, Tinnas D, Williams PM, Wood WL;
	XX	WPI; 2003-328860/31.
	DR	P-PsDB; ABU72207.
	XX	
	PT	New secreted and transmembrane nucleic acids and polypeptides, designated
	PT	as PRO, useful for treating inflammation, organ failure, atherosclerosis,
	PT	cardiac injury, infertility, birth defects, premature aging, AIDS, or
	PT	cancer.
	PS	Claim 2; Fig 38; 453pp; English.
	XX	
	CC	The invention describes an isolated nucleic acid (I) comprising, or which
	CC	is at least 80 % sequence identity to, or the full-length coding sequence
	CC	of, any of 118 300-2100 nucleotide sequences, which encodes its
	CC	corresponding PRO polypeptide selected from 118 100-700 amino acid
	CC	sequences, all given in the specification. The nucleic acids and
	CC	polypeptides are useful for treating inflammatory diseases, organ
	CC	failure, atherosclerosis, cardiac injury, infertility, birth defects,
	CC	premature aging, AIDS, cancer, or diabetic complications. The nucleic
	CC	acids are useful as hybridisation probes. In chromosome and gene mapping,
	CC	and in generating antisense RNA or DNA. The polypeptides are useful as
	CC	pharmaceuticals, diagnostics, biosensors or bioreactors. Both are useful
	CC	in tissue typing. This sequence encodes a novel human secreted and
	CC	transmembrane PRO polypeptide
	SQ	Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;
	Query Match	99.3%; Score 2387; DB 7; Length 2574;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 2387; Conservative	0; Mismatches 0; Indels 0; Gaps 0
QY	18	CTGCTCTCCGCCCGGTGTGAAGTGTGGGGGCGCTGGTGGAATGGCGTGTGCACCGA 77
Db	74	CTGCTCTCCGCCCGGTGTGAAGTGTGGGGGCGCTGGTGGAATGGCGTGTGCACCGA 133
QY	78	CGCGCGCTCCCTGTGAAGAAGAAAGTCTCAGCTAGAACAGAGCGGCTTAGTTTCGANAAG 137
Db	134	CGCGCGCTCCCTGTGAAGAAGAAAGTCTCAGCTAGAACAGAGCGGCTTAGTTTCGANAAG 193
QY	138	GAGGATCAGGGATGTTTTGCGAGCGGCTGGAAACAGAGCGTGCCTGATTAGAGAAAGCGGCT 197
Db	194	GAGGATCAGGGATGTTTTGCGAGCGGCTGGAAACAGAGCGTGCCTGATTAGAGAAAGCGGCT 253
QY	198	CAATGGCTCCCTCTCTGCTGTCAGCCCGCTGGCTGCTGTGTACTACGCTGTGCTCTTAAGC 257
Db	254	CAATGGCTCCCTCTCTGCTGTCAGCCCGCTGGCTGCTGTGTACTACGCTGTGCTCTTAAGC 313
QY	258	TACACCTCTGCGCGAGTTGGCTTGAGCTTCGCGGAACTTGGCTTTGCGGTGCGAGCTC 317
Db	314	TACACCTCTGCGCGAGTTGGCTTGAGCTTCGCGCGAACTTGGCTTTGCGGTGCGAGCTC 373
QY	318	TGTGCTGCAAAAAGGGCTCTTGAGACTGCGCGCTGGCGCGGCTGCGCGCAACCGGAAG 377
Db	374	TGTGCTGCAAAAAGGGCTCTTGAGACTGCGCGCTGGCGCGGCTGCGCGCAACCGGAAG 433
QY	378	GTCGCCAGGGGGGCTGACAGCTGTGCTGTGCGCTTGTGCGAATCTGSCCAAGCGCGCGC 437
Db	434	GTCGCCAGGGGGGCTGACAGCTGTGCTTGTGCGCTTGTGCGAATCTGSCCAAGCGCGCGC 493
QY	438	CGCACACCTTTTCTCATTTAAGGGCTTGGGGGCTTTTAGCTACTAGAGGCGAGACGCTGAGA 497
Db	494	CGCACACCTTTTCTCATTTAAGGGCTTGGGGGCTTTTAGCTACTAGAGGCGAGACGCTGAGA 553
QY	498	GTTAAAGAGGCTGTGCAAGCGCTTCTCTAAGTGTGCTAGAGCTGTGGAATCCGGAACGGCG 557
Db	554	GTTAAAGAGGCTGTGCAAGCGCTTCTCTAAGTGTGCTAGAGCTGTGGAATCCGGAACGGCG 613
QY	558	GCGACAACGCGCGAGGGAGCGCTGTGAGAAAGCAAGCGGCGACGCGCGGAGCCGGAATG 617

Db 614 GCGACAGCGGAGGAGCGCTGTGAGAAAGCGAGCGGCGAGCCCGGAGCCCGAGATG 673
Qy 618 CAGCGCGCGAAGCGCGCGAGATTTCGCGAGGAGCGATGCCGCCAGAGGTGAGAG 677
Db 674 CAGCGCGCGAAGCGCGCGAGATTTCGCGAGGAGCGATGCCGCCAGAGGTGAGAG 733
Qy 678 CCGCGCGCGCTGTGACACTGAGAGAACTGTGTGGCTCTCTCCCGCTGCGCGAGAGT 737
Db 734 CCGCGCGCGCTGTGACACTGAGAGAACTGTGTGGCTCTCTCCCGCTGCGCGAGAGT 793
Qy 738 TTCTGTGCTGTGCTGCGGCTGCGAGAGCGCGCTGCGAGTCTCTTGTGCCACG 797
Db 754 TTCTGTGCTGTGCTGCGGCTGCGAGAGCGCGCTGCGAGTCTCTTGTGCCACG 853
Qy 758 CCTGCGCGCGCGCGCGCTGCTGCACTGCTCCGAGCTGCGGCGCGCGCGCTGCTG 857
Db 854 CCTGCGCGCGCGCGCGCGCTGCTGCACTGCTCCGAGCTGCGGCGCGCGCGCTGCTG 913
Qy 858 TGGCGCGAGATTTCGAGAGTCCCTGAGAGCGCGAGCTGCGCGCGCTGAGAGCGAG 917
Db 914 TGGCGCGAGATTTCGAGAGTCCCTGAGAGCGCGAGCTGCGCGCGCTGAGAGCGAG 973
Qy 918 TCCACCTGTGGGCTGCGAGCGCGCGAGAACCCACTGCTGGAATTAGCGATTGCTGCTG 977
Db 974 TCCACCTGTGGGCTGCGAGCGCGCGAGAACCCACTGCTGGAATTAGCGATTGCTGCTG 1033
Qy 978 AAGTGTGCGCTGGAATGAGTGGGCGAGTGGCGAGTACCTCTCTTCCCGCGAGCATTA 1037
Db 1034 AAGTGTGCGCTGGAATGAGTGGGCGAGTGGCGAGTACCTCTCTTCCCGCGAGCATTA 1093
Qy 1038 CAGACAGCTGCTGTACACTCTTCACTGTGCGACCGCGCTCCCGAGAGCTGTGCGA 1097
Db 1094 CAGACAGCTGCTGTACACTCTTCACTGTGCGACCGCGCTCCCGAGAGCTGTGCGA 1153
Qy 1098 TCAGTCACTGTAAGATCTCTGCAATGCGAGGCTTTATACAGTGTGTGTGTCCACGAG 1157
Db 1154 TCAGTCACTGTAAGATCTCTGCAATGCGAGGCTTTATACAGTGTGTGTGTCCACGAG 1213
Qy 1158 AAGTGTGATCTACCTGCGCTCCCACTTACCAATGCTGCTGCTGCGGAGATG 1217
Db 1214 AAGTGTGATCTACCTGCGCTCCCACTTACCAATGCTGCTGCTGCGGAGATG 1273
Qy 1218 TGGGCTGCAATGCGGCAATTGGGCGCAAGTGTGTGAATTCGAAGTCTGCGCTGCTG 1277
Db 1274 TGGGCTGCAATGCGGCAATTGGGCGCAAGTGTGTGAATTCGAAGTCTGCGCTGCTG 1333
Qy 1278 TCTGGAGAGATGCGAGAGCAAGGAGTGTGCTGCAAGTGTGGGAGGCTGAGC 1337
Db 1334 TCTGGAGAGATGCGAGAGCAAGGAGTGTGCTGCAAGTGTGGGAGGCTGAGC 1393
Qy 1338 GATACCTGTCTACCAAGCCCGGAGCAAGGCAAGTGTGCCATTAAGTTCGGCTGCGAG 1397
Db 1394 GATACCTGTCTACCAAGCCCGGAGCAAGGCAAGTGTGCCATTAAGTTCGGCTGCGAG 1453
Qy 1398 TGGGAGAGGCGCTGCGGCGCAATACCTGAGAGCTTTGTGTGGGCGCTTGGGCGCGCTG 1457
Db 1454 TGGGAGAGGCGCTGCGGCGCAATACCTGAGAGCTTTGTGTGGGCGCTTGGGCGCGCTG 1513
Qy 1458 AGGTCTGAGACATATGAGTGAAGAGGAGGCAAGTGTGCCATTAAGTTCAGACAGAGC 1517
Db 1514 AGGTCTGAGACATATGAGTGAAGAGGAGGCAAGTGTGCCATTAAGTTCAGACAGAGC 1573
Qy 1518 AGCGGCGCGCTGTGGGCGCTGCTCTCTGCTTTAAGAGATCTTCCCTCTCTCTGA 1577
Db 1574 AGCGGCGCGCTGTGGGCGCTGCTCTCTGCTTTAAGAGATCTTCCCTCTCTCTGA 1633
Qy 1578 TTGCTATATGATCAACAGAGAGAGCAATTCGAGACCCCGAGGAGAGCTATAGGCA 1637
Db 1634 TTGCTATATGATCAACAGAGAGAGCAATTCGAGACCCCGAGGAGAGCTATAGGCA 1693
Qy 1638 CATCTCAGGTAGCCAGGCGCTGTGTGTGCGCGCGTAAAGCCAGAGTCCCATTCCTG 1697

Db 1694 CATCTCAGGTAGCCAGGCGCTGTGTGTGCGCGCGTAAAGCCAGAGTCCCATTCCTG 1753
Qy 1698 GCTATGCTGCGCGGCGCAAGCTGTGGCCAGGAGAGTGTCTTAAGATGTCTTCGCGCTG 1757
Db 1754 GCTATGCTGCGCGGCGCAAGCTGTGGCCAGGAGAGTGTCTTAAGATGTCTTCGCGCTG 1813
Qy 1758 GGGATGTTTCTTCAACACTGAGGAGCCTGTGTGTGCGAGTGAACAGAGTTTCTGCGCT 1817
Db 1814 GGGATGTTTCTTCAACACTGAGGAGCCTGTGTGTGCGAGTGAACAGAGTTTCTGCGCT 1873
Qy 1818 TCCATGATGCTACTGAGAGACCTTCAAGTGTGAAGGAGGAGAAATGTGCCACACCGAG 1877
Db 1874 TCCATGATGCTACTGAGAGACCTTCAAGTGTGAAGGAGGAGAAATGTGCCACACCGAG 1933
Qy 1878 TGGCAGAGGCTTCGAGGCGCTGATTTTCTTCAAGAGGAGAAAGTGTGAGTCACTG 1937
Db 1934 TGGCAGAGGCTTCGAGGCGCTGATTTTCTTCAAGAGGAGAAAGTGTGAGTCACTG 1993
Qy 1938 TGGCAGAGGCTTCGAGGCGCTGATTTTCTTCAAGAGGAGAAAGTGTGAGTCACTG 1997
Db 1994 TGGCAGAGGCTTCGAGGCGCTGATTTTCTTCAAGAGGAGAAAGTGTGAGTCACTG 2053
Qy 1998 TGGACCTTATGAGAGCTTACACCCAGTGTGTGAGAACTTGGCCACTTATGCCCGCGCC 2057
Db 2054 TGGACCTTATGAGAGCTTACACCCAGTGTGTGAGAACTTGGCCACTTATGCCCGCGCC 2113
Qy 2058 GATTCTCAGGCTCCAGAGAGTCTTTGGCCACACAGAGACCTTCAAGAGAGAAAGTTC 2117
Db 2114 GATTCTCAGGCTCCAGAGAGTCTTTGGCCACACAGAGACCTTCAAGAGAGAAAGTTC 2173
Qy 2118 GATGTGAATGAGGCGTTGAGACCCAGACCTGTGTGACCCACTGTACTGTCTGAGC 2177
Db 2174 GATGTGAATGAGGCGTTGAGACCCAGACCTGTGTGACCCACTGTACTGTCTGAGC 2233
Qy 2178 AGGCTGTAGGTGCTACCTGCGCTCCCTCAACAATGCGCGGTGACAGCGCGCTCTGAGAA 2237
Db 2234 AGGCTGTAGGTGCTACCTGCGCTCCCTCAACAATGCGCGGTGACAGCGCGCTCTGAGAA 2293
Qy 2238 ACCTTGAATCTGAGAACTTCAACACTGAGAGGCACTGTGAGAGAACTGTGTGGGCTG 2297
Db 2294 ACCTTGAATCTGAGAACTTCAACACTGAGAGGCACTGTGAGAGAACTGTGTGGGCTG 2353
Qy 2298 GGGCGCTGAGGTGATCTGCGGCTGTGAGGAGCTTTCTATACCAAGTGTGCGCTACT 2357
Db 2354 GGGCGCTGAGGTGATCTGCGGCTGTGAGGAGCTTTCTATACCAAGTGTGCGCTACT 2413
Qy 2358 ATTTTGAATTAATGTGCTGAGTGAATCCAGCTGTCTGACCTA 2404
Db 2414 ATTTTGAATTAATGTGCTGAGTGAATCCAGCTGTCTGACCTA 2460

RESULT 10
ACA71709
ID ACA71709 standard; cDNA; 2574 BP.
XX
XX ACA71709;
DT 11-AUG-2003 (first entry)
XX
XX Human secreted and transmembrane polypeptide PRO703 cDNA.
XX
XX Human; ss; gene; thrombolytic agent; interferon; interleukin; cytokine;
XX erythropoietin; colony stimulating factor; cancer; colorectal carcinoma;
XX apoptosis related condition; AIDS; amyotrophic lateral sclerosis;
XX inflammatory disease; asthma; atherosclerosis; neurodegenerative disease;
XX gastrointestinal disorder; Alzheimer's disease; Parkinson's disease;
XX hypertension; myocardial ischemia; kidney disease; carcinogenesis;
XX glomerulonephritis; lung disease; pulmonary hypertension; pre-eclampsia;
XX bronchial asthma; gastric ulcer; renal failure; cardiovascular disease;
XX inflammatory bowel disease; reproductive disorder; premature labour.
XX
XX Homo sapiens.
XX

```

PN US2002177553-A1.
XX
XX 28-NOV-2002.
PF 15-OCT-2001; 2001US-00978192.
XX
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-006511P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-0004022P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079636P.
PR 27-MAR-1998; 98US-0079636P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079785P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 26-JUN-1998; 98US-0010541P.
PR 07-OCT-1998; 98US-0016897P.
PR 07-OCT-1998; 98US-0016897P.
PR 06-NOV-1998; 98US-0018421P.
PR 06-NOV-1998; 98US-0018736P.
PR 20-NOV-1998; 98US-0024855P.
PR 07-DEC-1998; 98US-0020205P.
PR 22-DEC-1998; 98US-0021851P.
PR 05-JAN-1999; 99US-0000106P.
PR 08-MAR-1999; 99US-0025445P.
PR 10-MAR-1999; 99US-0025568P.
PR 10-MAR-1999; 99US-0025568P.
PR 12-MAR-1999; 99US-0026721P.
PR 12-APR-1999; 99US-0028429P.
PR 14-MAY-1999; 99US-0031183P.
PR 14-MAY-1999; 99US-0031073P.
PR 02-JUN-1999; 99US-0031225P.
PR 25-AUG-1999; 99US-0038013P.
PR 25-AUG-1999; 99US-0038013P.
PR 25-AUG-1999; 99US-0038013P.
PR 25-AUG-1999; 99US-0038013P.
PR 30-NOV-1999; 99US-0038013P.
PR 02-DEC-1999; 99US-0038013P.
PR 16-DEC-1999; 99US-0038013P.
PR 30-DEC-1999; 99US-0038013P.
PR 03-DEC-1999; 99US-0038013P.
PR 05-JAN-2000; 2000US-0000212P.
PR 06-JAN-2000; 2000US-0000277P.
PR 06-JAN-2000; 2000US-0000376P.
PR 11-FEB-2000; 2000US-0003365P.
PR 18-FEB-2000; 2000US-0004341P.
PR 24-FEB-2000; 2000US-0005004P.
PR 02-MAR-2000; 2000US-0005841P.
PR 10-MAR-2000; 2000US-0006311P.
PR 21-MAR-2000; 2000US-0007533P.
PR 30-MAR-2000; 2000US-0008439P.
PR 17-MAY-2000; 2000US-0013705P.
PR 22-MAY-2000; 2000US-0014042P.
PR 30-MAY-2000; 2000US-0014941P.
PR 02-JUN-2000; 2000US-0015264P.
PR 28-JUL-2000; 2000US-0020710P.
PR 24-AUG-2000; 2000US-0023328P.
PR 08-NOV-2000; 2000US-00709238.
PR 27-NOV-2000; 2000US-00723749.
PR 01-DEC-2000; 2000US-00732678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000US-00747259.
PR 28-FEB-2001; 2001US-00815744.
PR 22-MAR-2001; 2001US-00815920.
PR 22-MAR-2001; 2001US-00815920.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-00817092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001US-00871800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001US-0019692P.
PR 20-JUN-2001; 2001US-0021066P.
PR 09-JUL-2001; 2001US-0021735P.
PR 30-JUL-2001; 2001US-00918585.
XX
XX (GENTECH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
XX Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gertsen ME;
XX Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
XX Kijavrin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;
XX Stewart TA, Thomas D, Williams PM, Wood WI;
XX
XX WPI; 2003-328499/31.
XX
XX P-PSDB; AB094887.
XX
XX New isolated PRO polypeptides e.g. PRO213, PRO274 and PRO300, for use as
XX pharmaceuticals, diagnostics, biosensors and bioreactors, for identifying
XX modulators of receptor-ligand interactions.
XX
XX Claim 2; SEQ ID NO 101; 55pp; English.
XX
XX The invention relates to an isolated secreted and transmembrane
XX polypeptide, designated as PRO polypeptide. The PRO polypeptide is useful
XX in PRO polypeptide detection methods. The PRO polypeptide is useful for
XX linking a bioactive molecule to a cell. The PRO polypeptide or an
XX antibody against it is useful for modulating a biological activity of a
XX cell. The PRO polypeptide is useful in industrial applications including
XX pharmaceuticals, diagnostics, biosensors and bioreactors. The PRO
XX polypeptide is also useful as a thrombolytic agent, interferon,
XX interleukin, erythropoietin, colony stimulating factor and other
XX cytokines. The PRO polypeptide is useful for treating disease such as
XX cancer e.g. colorectal carcinoma; apoptosis related conditions e.g. AIDS,
XX amyotrophic lateral sclerosis; inflammatory disease e.g. asthma,
XX atherosclerosis; neurodegenerative disease e.g. Alzheimer's disease,
XX Parkinson's disease; cardiovascular disease e.g. hypertension and
XX myocardial ischaemia; kidney disease e.g. renal failure and
XX glomerulonephritis; lung disease e.g. pulmonary hypertension, bronchial
XX asthma; gastrointestinal disorders e.g. gastric ulcer and inflammatory
XX bowel disease; reproductive disorders e.g. premature labour and
XX pre-eclampsia; carcinogenesis. The present sequence represents a cDNA
XX encoding a PRO polypeptide of the invention. Note: The sequence data for
XX this patent did not form part of the printed specification but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?docid=20020177553
XX
XX Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;
SQ
Query Match 99.3%; Score 2387; DB 7; Length 2574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY 18 CTGCTCTCCGCGCCGCTGGAGTGGAGGCGCTGGAGTGGAGTGGAGCGGCTCCAGCGCA 77
Db 74 CTGCTCTCCGCGCCGCTGGAGTGGAGGCGCTGGAGTGGAGTGGAGCGGCTCCAGCGCA 133

```

QY 78 CGCGGCTCTCTGAGAGAGAGTCTCAGCTAGAAAGAGCGGCGCTAGGTTTCGGAGG 137
Db 134 CGCGGCTCTCTGAGAGAGAGTCTCAGCTAGAAAGAGCGGCGCTAGGTTTCGGAGG 193
QY 138 GAGGATCAGGAGTGTTCGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 197
Db 194 GAGGATCAGGAGTGTTCGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 253
QY 198 CGATGCTGCGGCTCTCTGAGAGAGTGTTCGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 257
Db 254 CGATGCTGCGGCTCTCTGAGAGAGTGTTCGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 313
QY 258 TACACCTCTGAGCGGCTGAGAGTGTTCGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 317
Db 314 TACACCTCTGAGCGGCTGAGAGTGTTCGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 373
QY 318 TGTGCTGCAAAAGGCTCTTCAGAGCTCGGCGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 377
Db 374 TGTGCTGCAAAAGGCTCTTCAGAGCTCGGCGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 433
QY 378 GTCCGAGGCGGCTGAGAGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 437
Db 434 GTCCGAGGCGGCTGAGAGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 493
QY 438 CGCAGCCTTTCTCATTCAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 497
Db 494 CGCAGCCTTTCTCATTCAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 553
QY 498 GTTACAGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 557
Db 554 GTTACAGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 613
QY 558 GCGACAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 617
Db 614 GCGACAGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 673
QY 618 CAGCGGCTGAGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 677
Db 674 CAGCGGCTGAGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 733
QY 678 CGCGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 737
Db 734 CGCGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 793
QY 738 TTTGCTGCTCTGAGTGTTCGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 797
Db 794 TTTGCTGCTCTGAGTGTTCGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 853
QY 798 CCTGCGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 857
Db 854 CCTGCGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 913
QY 858 TGGCGCGAGAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 917
Db 914 TGGCGCGAGAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 973
QY 918 TCCACCTGAGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 977
Db 974 TCCACCTGAGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1033
QY 978 AAGTGTGAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1037
Db 1034 AAGTGTGAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1093
QY 1038 CAGACAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1097
Db 1094 CAGACAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1153
QY 1098 TGAAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1157
Db 1154 TGAAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1213
QY 1158 AAGATGTATCTAAGTGTTCGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1217

Db 1214 AAGATGTATCTAAGTGTTCGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1273
QY 1218 TGGCTGTGATGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1277
Db 1274 TGGCTGTGATGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1333
QY 1278 TGTGGAAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1337
Db 1334 TGTGGAAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1393
QY 1338 GATACCTTTGACACGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1397
Db 1394 GATACCTTTGACACGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1453
QY 1398 TGGCGAGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1457
Db 1454 TGGCGAGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1513
QY 1458 AGTGTGAGAGATGAGAGCTGAGAGGCGAGAGCGGTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1517
Db 1514 AGTGTGAGAGATGAGAGCTGAGAGGCGAGAGCGGTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1573
QY 1518 AGCGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1577
Db 1574 AGCGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1633
QY 1578 TTTGCTGATGATGACACGAGAGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1637
Db 1634 TTTGCTGATGATGACACGAGAGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1693
QY 1638 CATCTCAGGTGAGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1697
Db 1694 CATCTCAGGTGAGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1753
QY 1698 GCTATGCTGAGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1757
Db 1754 GCTATGCTGAGGCGGCTGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1813
QY 1758 GGGATGTTTCTTCAACATGAGGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1817
Db 1814 GGGATGTTTCTTCAACATGAGGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1873
QY 1818 TCCATGATGATGAGAGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1877
Db 1874 TCCATGATGATGAGAGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1933
QY 1878 TGGCAGAGGCTTCAAGGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1937
Db 1934 TGGCAGAGGCTTCAAGGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1993
QY 1938 TGGCAGGAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 1997
Db 1994 TGGCAGGAGTGTTCGAGCGGCTGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 2053
QY 1998 TGGACCTTATGAGCGGCTTCAACGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 2057
Db 2054 TGGACCTTATGAGCGGCTTCAACGAGCGGCTGAGCGGCTGAGACGAGCGGTGCGATAGAGAGAGCGGCT 2113
QY 2058 GATTCCTCAGGCTTCAAGAGCGGCTTGGCGACACAGAGAGCGGCTTCAACAGAGAGAGCGGCT 2117
Db 2114 GATTCCTCAGGCTTCAAGAGCGGCTTGGCGACACAGAGAGCGGCTTCAACAGAGAGAGCGGCT 2173
QY 2118 GATTCGAAATGAGAGGCTTGAAGCGGCTTGGCGACACAGAGAGCGGCTTGAAGAGAGAGCGGCT 2177
Db 2174 GATTCGAAATGAGAGGCTTGAAGCGGCTTGGCGACACAGAGAGCGGCTTGAAGAGAGAGCGGCT 2233
QY 2178 AGGCTGAGGCTGAGCGGCTTGGCGACACAGAGAGCGGCTTGAAGAGAGAGCGGCTTGGCGAGAG 2237
Db 2234 AGGCTGAGGCTGAGCGGCTTGGCGACACAGAGAGCGGCTTGAAGAGAGAGCGGCTTGGCGAGAG 2293
QY 2238 ACCTTGAAATCTGAGAGCTTCAACGAGGAGCGGCTTGAAGAGAGAGCGGCTTGAAGAGAGAGCGGCT 2297

Db 2294 ACCTTGGAATCGAGACTTCCACACCTAGGACCTGAGAGAGAACTCTGTGGGGTGG 2353
 QY 2298 GGGCCGCTGAGGCTGACTGGGCTGTCAGAGATCTTTTCTATACAGAACTGGGCTCACT 2357
 Db 2354 GGGCCGCTGAGGCTGACTGGGCTGTCAGAGATCTTTTCTATACAGAACTGGGCTCACT 2413
 QY 2358 ATTTTCTATTAATGCTGAGCTGAGCTGAGCACTGCTCTCTACCTA 2404
 Db 2414 ATTTTCTATTAATGCTGAGCTGAGCTGAGCTGAGCTGCTCTCTACCTA 2460

RESULT 11
 ABX92349
 ID ABX92349 standard; cDNA; 2574 BP.
 AC ABX92349;
 XX
 XX 08-MAY-2003 (first entry)
 XX
 DE cDNA encoding human PRO703 polypeptide.
 XX
 KW Human; PRO polypeptide; secreted and transmembrane protein;
 KW Immune disorder; diabetes; hyper-insulinaemia; hypo-insulinaemia;
 KW Cardiac insufficiency; nervous system disorder; kidney disorder;
 KW Bone disorder; cartilage disorder; arthritis; tumour; wound healing;
 KW Genetic disorder; cytostatic; antidiabetic; antiinflammatory;
 KW Antiarthritic; anti-tumour; vulnery; antianaemic; dermatological;
 KW Cardiant; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002169284-A1.
 XX
 PD 14-NOV-2002.
 XX
 PF 16-OCT-2001; 2001US-00978697.
 XX
 XX 26-MAY-1981; 81US-00267213.
 PR 17-OCT-1997; 97US-0062250P.
 PR 03-NOV-1997; 97US-0064249P.
 PR 13-NOV-1997; 97US-0065311P.
 PR 21-NOV-1997; 97US-0066364P.
 PR 10-MAR-1998; 98US-0077450P.
 PR 11-MAR-1998; 98US-0077632P.
 PR 11-MAR-1998; 98US-0077641P.
 PR 11-MAR-1998; 98US-0077649P.
 PR 12-MAR-1998; 98US-0077791P.
 PR 13-MAR-1998; 98US-0078004P.
 PR 17-MAR-1998; 98US-0004022P.
 PR 20-MAR-1998; 98US-0078886P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 25-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079294P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079689P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 27-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 26-JUN-1998; 98US-0010541P.
 PR 07-OCT-1998; 98US-0016897P.
 PR 07-OCT-1998; 98US-0016897P.
 PR 02-NOV-1998; 98US-0018421P.
 PR 06-NOV-1998; 98US-0018736P.
 PR 20-NOV-1998; 98US-0020248P.
 PR 07-DEC-1998; 98US-0020205P.
 PR 22-DEC-1998; 98US-0021851P.
 PR 05-JAN-1999; 98US-0020106P.
 PR 05-MAR-1999; 98US-0025346P.
 PR 08-MAR-1999; 98US-0025346P.
 PR 08-MAR-1999; 98US-0025346P.

PR 10-MAR-1999; 98US-00265686.
 PR 10-MAR-1999; 98US-00265686.
 PR 12-APR-1999; 98US-00284291.
 PR 14-MAY-1999; 98US-00311832.
 PR 14-MAY-1999; 98US-00311832.
 PR 02-JUN-1999; 98US-00311832.
 PR 25-AUG-1999; 98US-00380137.
 PR 25-AUG-1999; 98US-00380137.
 PR 25-AUG-1999; 98US-00380137.
 PR 25-AUG-1999; 98US-00380137.
 PR 25-AUG-1999; 98US-00380137.
 PR 30-DEC-1999; 98US-00380137.
 PR 30-DEC-1999; 98US-00380137.
 PR 05-JAN-2000; 98US-00380137.
 PR 06-JAN-2000; 98US-00380137.
 PR 06-JAN-2000; 98US-00380137.
 PR 11-FEB-2000; 98US-00380137.
 PR 18-FEB-2000; 98US-00380137.
 PR 24-FEB-2000; 98US-00380137.
 PR 02-MAR-2000; 98US-00380137.
 PR 10-MAR-2000; 98US-00380137.
 PR 21-MAR-2000; 98US-00380137.
 PR 30-MAR-2000; 98US-00380137.
 PR 17-MAY-2000; 98US-00380137.
 PR 22-MAY-2000; 98US-00380137.
 PR 30-MAY-2000; 98US-00380137.
 PR 02-JUN-2000; 98US-00380137.
 PR 28-JUL-2000; 98US-00380137.
 PR 28-AUG-2000; 98US-00380137.
 PR 08-NOV-2000; 98US-00380137.
 PR 27-NOV-2000; 98US-00380137.
 PR 01-DEC-2000; 98US-00380137.
 PR 20-DEC-2000; 98US-00380137.
 PR 28-FEB-2001; 98US-00380137.
 PR 28-FEB-2001; 98US-00380137.
 PR 22-MAR-2001; 98US-00380137.
 PR 22-MAR-2001; 98US-00380137.
 PR 22-MAR-2001; 98US-00380137.
 PR 10-MAY-2001; 98US-00380137.
 PR 25-MAY-2001; 98US-00380137.
 PR 01-JUN-2001; 98US-00380137.
 PR 01-JUN-2001; 98US-00380137.
 PR 05-JUN-2001; 98US-00380137.
 PR 14-JUN-2001; 98US-00380137.
 PR 19-JUN-2001; 98US-00380137.
 PR 20-JUN-2001; 98US-00380137.
 PR 29-JUN-2001; 98US-00380137.
 PR 09-JUL-2001; 98US-00380137.
 PR 30-JUL-2001; 98US-00380137.

(GENT) GENENTECH INC.
 XX
 PI Ashkenazi A, Baker KP, Borstein D, Desnoyers L, Eaton D;
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gertsen ME;
 PI Goddard A, Godowski PJ, Grimaldi JC, Gueney AL, Hillan KJ, Shelton DL;
 PI Kijavlin IJ, Kuo SS, Napier MA, Pan J, Paoi NF, Roy MA, Shelton DL;
 PI Stewart TA, Tumas D, Williams PM, Wood WI;
 XX
 DR WPI: 2003-288163/28.
 DR P-PSDB; AB01085.
 XX
 PT Novel secreted and transmembrane polypeptides and polynucleotides
 PT encoding them useful for treating cancer, kidney diseases, bone,
 PT cartilage disorders and immune deficiencies.
 XX
 PS Claim 2; Fig 38; 459pp; English.
 XX
 CC The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The PRO
 CC polypeptides are secreted and transmembrane proteins. The PRO


```
DB 1874 TCCATGATCGTACTGAGACACCTTCAGGTGGAAAGGGGAGATGTGGCCACAAACGAGG 1933
QY 1878 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 1937
DB 1934 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 1933
QY 1938 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 1937
DB 1994 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 2053
QY 1998 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 2057
DB 2054 TGGCAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 2113
QY 2058 GATTCCTAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 2117
DB 2114 GATTCCTAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 2173
QY 2118 GATTCCTAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 2177
DB 2174 GATTCCTAGAGGTCTTCAGAGCCCTAGATTTCTTCAGAGAGTAAAGCTATAGATCACTG 2233
QY 2178 AGGCTGATAGTGTCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTG 2237
DB 2234 AGGCTGATAGTGTCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTG 2293
QY 2238 AGGCTGATAGTGTCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTG 2297
DB 2294 AGGCTGATAGTGTCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTG 2353
QY 2298 AGGCTGATAGTGTCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTG 2357
DB 2354 AGGCTGATAGTGTCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTG 2413
QY 2358 AGGCTGATAGTGTCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTG 2404
DB 2414 AGGCTGATAGTGTCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTGCTTACCTG 2460

RESULT 12
ID ACA66090 standard; cDNA; 2574 BP.
AC ACA66090;
XX
AC ACA66090;
XX
DT 24-JUN-2003 (first entry)
XX
DE Human cDNA encoding secreted/transmembrane protein PRO703.
XX
KW Human; ss; gene; secreted protein; transmembrane protein; PRO;
KW malignancy; cancer; ovarian cancer; colorectal cancer; sarcoma;
KW leukemia; lymphoma; inflammatory disease; necrosis; atherosclerosis;
KW infertility; premature aging; psoriasis; inflammatory disease;
KW renal disease; arthritis; immune-mediated alopecia; stroke; encephalitis;
KW hepatitis; multiple sclerosis; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003004102-A1.
XX
PD 02-JAN-2003.
XX
PF 15-OCT-2001; 2001US-00978189.
XX
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-006511P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-0078022P.
PR 20-MAR-1998; 98US-0078086P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 25-MAR-1998; 98US-0079294P.
PR 25-MAR-1998; 98US-0079656P.
PR 27-MAR-1998; 98US-0079662P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 26-JUN-1998; 98US-0010541P.
PR 07-OCT-1998; 98US-0016897P.
PR 07-OCT-1998; 98US-0021141P.
PR 02-NOV-1998; 98US-0018421P.
PR 06-NOV-1998; 98US-0018736P.
PR 20-NOV-1998; 98US-0024855P.
PR 07-DEC-1998; 98US-0020054P.
PR 22-DEC-1998; 98US-0021851P.
PR 05-JAN-1999; 99US-0000106P.
PR 05-MAR-1999; 99US-0025446P.
PR 08-MAR-1999; 99US-0005028P.
PR 10-MAR-1999; 99US-0026568P.
PR 10-MAR-1999; 99US-0005190P.
PR 12-MAR-1999; 99US-0026721P.
PR 12-MAR-1999; 99US-0028429P.
PR 14-MAY-1999; 99US-0031832P.
PR 14-MAY-1999; 99US-0010733P.
PR 02-JUN-1999; 99US-0012252P.
PR 25-AUG-1999; 99US-0038013P.
PR 25-AUG-1999; 99US-0038013P.
PR 25-AUG-1999; 99US-0038014P.
PR 30-NOV-1999; 99US-0028313P.
PR 02-DEC-1999; 99US-0028551P.
PR 02-DEC-1999; 99US-0028551P.
PR 16-DEC-1999; 99US-0030095P.
PR 30-DEC-1999; 99US-0031243P.
PR 30-DEC-1999; 99US-0031274P.
PR 05-JAN-2000; 2000US-0000219P.
PR 06-JAN-2000; 2000US-0000277P.
PR 06-JAN-2000; 2000US-0000376P.
PR 11-FEB-2000; 2000US-0003565P.
PR 18-FEB-2000; 2000US-0004341P.
PR 24-FEB-2000; 2000US-0005004P.
PR 01-MAR-2000; 2000US-0005601P.
PR 02-MAR-2000; 2000US-0005841P.
PR 10-MAR-2000; 2000US-0005839P.
PR 21-MAR-2000; 2000US-0007532P.
PR 30-MAR-2000; 2000US-0008439P.
PR 17-MAY-2000; 2000US-0013705P.
PR 22-MAY-2000; 2000US-0014042P.
PR 30-MAY-2000; 2000US-0014941P.
PR 02-JUN-2000; 2000US-0015264P.
PR 28-JUN-2000; 2000US-0020271P.
PR 24-AUG-2000; 2000US-0023328P.
PR 08-NOV-2000; 2000US-00709238P.
PR 10-NOV-2000; 2000US-0030873P.
PR 27-NOV-2000; 2000US-00723749P.
PR 01-DEC-2000; 2000US-0032678P.
PR 20-DEC-2000; 2000US-00747259P.
PR 20-DEC-2000; 2000US-0034956P.
PR 28-FEB-2001; 2001US-0008520P.
PR 22-MAR-2001; 2001US-00816920P.
PR 22-MAR-2001; 2001US-00816920P.
PR 10-MAY-2001; 2001US-00854208P.
PR 10-MAY-2001; 2001US-00854280P.
PR 25-MAY-2001; 2001US-00854280P.
PR 01-JUN-2001; 2001US-00872035P.
```

PR 01-JUN-2001; 2001WO-US017800.
 PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 PR 30-JUL-2001; 2001US-00918585.
 (GENT) GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
 XX Ferrara N, Flivarov E, Fong S, Gao W, Gerber H, Gerritsen ME;
 XX Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
 XX Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;
 XX Stewart TA, Tunas D, Williams PM, Wood WJ;
 DR WPI, 2003-341189/32.
 DR P-PSDB; AB080354.
 XX New genes and secreted and transmembrane polypeptides (e.g. PRO337 or
 PT PRO159), useful for treating or diagnosing e.g. cancers,
 PT atherosclerosis, infertility, stroke, encephalitis, hepatitis or multiple
 PT sclerosis in mammals.
 XX Claim 2, Fig 38, 460pp; English.
 XX The invention relates to a new isolated nucleic acid molecule comprising a
 CC sequence with at least 80% identity to: (a) a nucleotide encoding any of
 CC 94 PRO polypeptides whose sequences are fully defined in the
 CC specification; or (b) any of 94 nucleotide sequences fully defined in the
 CC specification; or the full length coding sequence of any these 94
 CC nucleotide sequences. Also included are an isolated PRO polypeptide
 CC scoring at least 80% positives when compared to any of the PRO
 CC polypeptide sequences cited above (or an isolated PRO polypeptide having
 CC at least 80% amino acid sequence identity to: (a) an amino acid sequence
 CC encoded by the nucleotide deposited with ATCC numbers listed in the
 CC specification; (b) the PRO polypeptide, lacking its associated signal
 CC peptide; or (c) an extracellular domain of the PRO polypeptide, with or
 CC lacking its associated signal peptide), a vector comprising the nucleic
 CC acid molecule, a host cell comprising the vector (and producing a PRO
 CC polypeptide), a chimeric molecule comprising the PRO polypeptide fused
 CC to a heterologous amino acid sequence and an anti-PRO antibody. The PRO
 CC polypeptides or polynucleotides are useful as pharmaceuticals,
 CC diagnostics, biosensors or bioreactors. These are particularly useful for
 CC detecting or treating e.g. malignancies or cancers (e.g. ovarian cancer,
 CC colorectal cancer, sarcoma, leukaemia or lymphoma), inflammatory disease,
 CC necrosis, atherosclerosis, infertility, premature aging, psoriasis,
 CC inflammatory disease, renal disease, arthritis, immune-mediated alopecia,
 CC stroke, encephalitis, hepatitis, or multiple sclerosis in mammals. The
 CC PRO polypeptides are useful in drug screening, particularly as targets
 CC for therapeutic intervention in these diseases, and in the diagnostic
 CC determination of the presence of these diseases. The PRO polypeptides are
 CC also useful as molecular weight markers, or for chromosome
 CC identification. The PRO genes are useful as hybridisation probes, or for
 CC screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may
 CC also be used in gene therapy, particularly for replacing a defective
 CC gene. The present sequence encodes a PRO polypeptide
 XX
 XX Sequence 2574 BP; 470 A; 775 C; 821 G; 508 T; 0 U; 0 Other;
 SQ
 Query Match 99.3%; Score 2387; DB 7; Length 2574;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 GAGATCAGGAGATGTTTGGAGAGCGGCTGGAACAGACGGTGCATAGAGAAAGCGGCT 197
 Db 194 GAGATCAGGAGATGTTTGGAGAGCGGCTGGAACAGACGGTGCATAGAGAAAGCGGCT 253
 QY 198 CCAATGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 257
 Db 254 CCAATGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 313
 QY 258 TACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 317
 Db 314 TACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 373
 QY 318 TGTGCTGCAAAAGAGGCTCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377
 Db 374 TGTGCTGCAAAAGAGGCTCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 433
 QY 378 GTCCGAGGAGGAGGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 437
 Db 434 GTCCGAGGAGGAGGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
 QY 438 CGCACACCTTTCTCATTTACAGGCTCGCGGCTTTAGCTACTAGAGAGCGGAGCGGAGA 497
 Db 494 CGCACACCTTTCTCATTTACAGGCTCGCGGCTTTAGCTACTAGAGAGCGGAGCGGAGA 553
 QY 498 GTAAACAGGCTGACAGCGCTTCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 557
 Db 554 GTAAACAGGCTGACAGCGCTTCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 613
 QY 558 GGCACAGCGGCGAGGAGGAGGCTGAGAGAGCGAGGAGCGGAGCGGAGCGGAGAGT 617
 Db 614 GGCACAGCGGCGAGGAGGAGGCTGAGAGAGCGAGGAGCGGAGCGGAGCGGAGAGT 673
 QY 618 CAGCGGCGGAGAGCGGAGGAGGCTGAGAGAGCGAGGAGCGGAGCGGAGAGT 677
 Db 674 CAGCGGCGGAGAGCGGAGGAGGCTGAGAGAGCGAGGAGCGGAGCGGAGAGT 733
 QY 678 CGCGCGCGCTCTGTACCTGAGCAACTGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 737
 Db 734 CGCGCGCGCTCTGTACCTGAGCAACTGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
 QY 738 TTCTGTGCTCTGTGTTGAGGCTGAGGAGCGGAGCGGAGCGGAGCGGAGAGT 797
 Db 794 TTCTGTGCTCTGTGTTGAGGCTGAGGAGCGGAGCGGAGCGGAGCGGAGAGT 853
 QY 798 CCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 857
 Db 854 CCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 913
 QY 858 TGGCGCAGAGTTTCTGAGAGTCCCTGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 917
 Db 914 TGGCGCAGAGTTTCTGAGAGTCCCTGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 973
 QY 918 TCCACCTGTGAGGCTGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 977
 Db 974 TCCACCTGTGAGGCTGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1033
 QY 978 AAGTGTCCGCTGAGAGTGGGCTGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1037
 Db 1034 AAGTGTCCGCTGAGAGTGGGCTGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1093
 QY 1038 CAGACAGGAGCTGTACATCTTCACTCTGAGCAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1097
 Db 1094 CAGACAGGAGCTGTACATCTTCACTCTGAGCAAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1153
 QY 1098 TCAGTATCTGAGAGTCTGAGAGTCCGAGGAGCTTATCAGAGTGTGAGTTCACACAG 1157
 Db 1154 TCAGTATCTGAGAGTCTGAGAGTCCGAGGAGCTTATCAGAGTGTGAGTTCACACAG 1213
 QY 1158 AAGATGATCTACCTGCGCTTCCACTACCAATGTCGAGTTCCTGCTGAGGATCG 1217
 Db 1214 AAGATGATCTACCTGCGCTTCCACTACCAATGTCGAGTTCCTGCTGAGGATCG 1273
 QY 1218 TGGGCTGATAGGAGATTTGGGCGCACAGTGTGTGTAATTCAGATTCTGCGCTGCTCACT 1277

D	1274	TGGGCTGCATGGGCACTTGGGGGCAACAAGTGGCTGAATTCAGATCTCGGGCTGGCACT	1333
Q	1278	TCGAGAAATTCGACAGCAACAAGGTGACGGTGTTCAGTACATTGGGAGCTGTGCC	1337
D	1334	TCGGGAAATTCGACAGCAACAAGGTGACGGTGTTCAGTACATTGGGAGCTGTGCC	1393
Q	1338	GATACCTTCAACCAAGGCCCCGAGCAAGGAGAAAGTGGCAATTAAGTTCGGGCTGGAGC	1397
D	1394	GATACCTTGTCAACCAAGGCCCCGAGCAAGGAGAAAGTGGCAATTAAGTTCGGGCTGGAGC	1453
Q	1398	TGGGACGCGGGCTGGCCCAATACCTGGAGCCGTTTGTGGGCGCTTGGGCGCCCTG	1457
D	1454	TGGGACGCGGGCTGGCCCAATACCTGGAGCCGTTTGTGGGCGCTTGGGCGCCCTG	1513
Q	1458	AGGTGCTGGAGCAATATGACCTGACAGGGCAACGGGGACCAATCACTACACAGAC	1517
D	1514	AGGTGCTGGAGCAATATGACCTGACAGGGCAACGGGGACCAATCACTACACAGAC	1573
Q	1518	AGGGGGGCGCTGTGGGGGGTGGCTTCCTGGCTTTACAGACATCTTCCGCTTCCTTGA	1577
D	1574	AGGGGGGCGCTGTGGGGGGTGGCTTCCTGGCTTTACAGACATCTTCCGCTTCCTTGA	1633
Q	1578	TTGGCTATATATCAACACAGAGAGCCAAATTGGGACCCCCAGGGGCACTGATATGGCA	1637
D	1634	TTGGCTATATATCAACACAGAGAGCCAAATTGGGACCCCCAGGGGCACTGATATGGCA	1693
Q	1638	CACTCTCAGGTAGACAGAGGCTGCTGGTGCCCGGTAACACAGACATCCCATTCCTGG	1697
D	1694	CATCTCAGGTAGACAGAGGCTGCTGGTGCCCGGTAACACAGACATCCCATTCCTGG	1753
Q	1698	GCTATGCTGGCGGGCCAGAGCTGGCCAGGGGAAATTGCTAAAGATGCTTTCGGGCTG	1757
D	1754	GCTATGCTGGCGGGCCAGAGCTGGCCAGGGGAAATTGCTAAAGATGCTTTCGGGCTG	1813
Q	1758	GGGATTTTCTTCAACACTGGGAGCCGCGTGGCTGACATGACCAAGGTTTTCTCGCT	1817
D	1814	GGGATTTTCTTCAACACTGGGAGCCGCGTGGCTGACATGACCAAGGTTTTCTCGCT	1873
Q	1818	TCCATGATGTACTGAGACACTTCAGGTGAAAGGGGAGAAATGTGGCCACACCGAG	1877
D	1874	TCCATGATGTACTGAGACACTTCAGGTGAAAGGGGAGAAATGTGGCCACACCGAG	1933
Q	1878	TGGCAGAAGTCTTCAAGGCCCTAATTTCTTCAAGAGGTGAACGCTATATGAGATCACTG	1937
D	1934	TGGCAGAAGTCTTCAAGGCCCTAATTTCTTCAAGAGGTGAACGCTATATGAGATCACTG	1993
Q	1938	TGGCAGGCGATTAAGAGGAGGCTGGAATGGCAGCCTAGTTCGCGTCCGCCCAAGCTT	1997
D	1994	TGGCAGGCGATTAAGAGGAGGCTGGAATGGCAGCCTAGTTCGCGTCCGCCCAAGCTT	2053
Q	1998	TGGACTTATGACGCTCTACACCACTGTGTGAGAACTTGCACTTATGCCCGGCCCC	2057
D	2054	TGGACTTATGACGCTCTACACCACTGTGTGAGAACTTGCACTTATGCCCGGCCCC	2113
Q	2058	GATTCCTCAGGCTCCAGGAGTCTTTGGGCAACAAGAGACCTTCAACAGCAGAAAGTTTC	2117
D	2114	GATTCCTCAGGCTCCAGGAGTCTTTGGGCAACAAGAGACCTTCAACAGCAGAAAGTTTC	2173
Q	2118	GGATGGCAATAGAGGCTTCAACCCCAAGACCTGTGTGACCACTGTATCGTATCTGAAAC	2177
D	2174	GGATGGCAATAGAGGCTTCAACCCCAAGACCTGTGTGACCACTGTATCGTATCTGAAAC	2233
Q	2178	AGGCTGTAGATGCCCAACCGGCCCTCAACAATGGCCGGTACAGGCGCCCTGGGACAGAA	2237
D	2234	AGGCTGTAGATGCCCAACCGGCCCTCAACAATGGCCGGTACAGGCGCCCTGGGACAGAA	2293
Q	2238	ACCTTGAAATCGAAGACTTCAACACTTGAAGGCACTGAGAAGGAATCTGTGGGGATGG	2297
D	2294	ACCTTGAAATCGAAGACTTCAACACTTGAAGGCACTGAGAAGGAATCTGTGGGGATGG	2353
Q	2298	GGGCGGTGCAAGTGTACTGGGCTGTCAAGGAATTTTCTATACAGAACTGCGGTCACT	2357

```
PR 15-APR-1998; 98US-0081955P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083336P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083554P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084419P.
PR 07-MAY-1998; 98US-0084598P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 13-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085573P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085689P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086022P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087088P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98WO-US021141.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98WO-US024855.
PR 22-DEC-1998; 98US-0113296P.
PR 23-DEC-1998; 98US-0113621P.
PR 05-JAN-1999; 99WO-US000106.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005190.
PR 12-MAR-1999; 99US-0123957P.
PR 12-MAR-1999; 99US-0126773P.
PR 29-MAR-1999; 99US-0130232P.
PR 21-APR-1999; 99US-0131022P.
PR 26-APR-1999; 99US-0131445P.
PR 14-MAY-1999; 99US-0134287P.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 16-JUN-1999; 99US-0139557P.
PR 23-JUN-1999; 99US-0141037P.
PR 07-JUL-1999; 99US-0142680P.

PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 28-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 03-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US000376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 24-FEB-2000; 2000WO-US005004.
PR 24-FEB-2000; 2000WO-US005041.
PR 10-MAR-2000; 2000WO-US005841.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001WO-US009552.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001WO-US017800.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 30-JUL-2001; 2001US-00918585.

PR PA (GETH ) GENENTECH INC.
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Perrana N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ.
PI Kijavitt IU, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;
PI Stewart TA, Tumas D, Williams PM, Wood WI;
XX WPI; 2003-521814/49.
XX P-PSDB; ADA24641.
XX DR
XX PT New isolated PRO polypeptides for example extracellular, secreted and
XX PT membrane bound proteins, useful for modulating the biological activities
XX PT of cells and for treating, for example diabetes, cancer, rheumatoid
XX PT arthritis, and hearing loss.
XX Claim 2; Fig 38; 461p; English.
XX XX
CC The invention describes an isolated secreted and transmembrane (PRO)
CC polypeptide (I). PRO337 polypeptide is useful for detecting PRO493
CC polypeptide in a sample, and vice versa. PRO725, PRO700 and PRO739 are
CC useful for detecting PRO1559 polypeptide in a sample, and PRO1559 is
CC useful for detecting PRO725, PRO700 and PRO739 in a sample. PRO493 is
CC useful for linking a bioactive molecule to a cell expressing a PRO337
CC polypeptide, and PRO337 is useful for linking a bioactive molecule to a
CC cell expressing a PRO493 polypeptide. PRO1559 is useful for linking a
CC bioactive molecule to a cell expressing a PRO735, PRO700 and PRO739
CC CC

Query Match 99.3%; Score 2387; DB 8; Length 2574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 CTGCTCTCCGCGCCGCTGAGTGTGGGGCCGCTGGAGTGGCGTGCAGCGCA 77
Db 74 CTGCTCTCCGCGCCGCTGAGTGTGGGGCCGCTGGAGTGGCGTGCAGCGCA 133
QY 78 CGCGCGCTCCCTCGAAGAGACTCTACCTAGAACGAGCGGCCCTTAGCTTTGCGAAG 137
```


QY 1158 AAGATGTGATCTACCTCGCCCTCCCACTCTACACATGTCGSGTTCCCTGTGCGGATCG 1217
Db 1214 AAGATGTGATCTACCTCGCCCTCCCACTCTACACATGTCGSGTTCCCTGTGCGGATCG 1273
QY 1218 TGGGCTGCATGGGCAATGGGGCCCAAGTGTGTAATCCAAATTCGAGTTCTCGCTGGTCACT 1277
Db 1274 TGGGCTGCATGGGCAATGGGGCCCAAGTGTGTAATCCAAATTCGAGTTCTCGCTGGTCACT 1333
QY 1278 TCTGGGAAGATTGGCAGACAGCAGGAGTGAAGGTTCTCCAGTACATTTGGGAGAGCTGTGCC 1337
Db 1334 TCTGGGAAGATTGGCAGACAGGAGTGAAGGTTCTCCAGTACATTTGGGAGAGCTGTGCC 1393
QY 1338 GATACCTTGTCAACCAAGCCCGGAGACAGAAAGTGGCCATTAAGGTCGCGCTGCGAG 1397
Db 1394 GATACCTTGTCAACCAAGCCCGGAGACAGAAAGTGGCCATTAAGGTCGCGCTGCGAG 1453
QY 1398 TGGGAGCGGCGCTGGGCGCCAGATACCTGGAGAGCTTTGTGGGCGCTTGGGCGCCCTGC 1457
Db 1454 TGGGAGCGGCGCTGGGCGCCAGATACCTGGAGAGCTTTGTGGGCGCTTGGGCGCCCTGC 1513
QY 1458 AGGTCTGAGACATATGACTGACAGAGGCGCAAGTGGCCACATCACTACACAGAGAC 1517
Db 1514 AGGTCTGAGACATATGACTGACAGAGGCGCAAGTGGCCACATCACTACACAGAGAC 1573
QY 1518 AGCGGGGCGCTGTGGGCGCTGCTTCTGCTTACAGACATATCTTCCCTTCTCTTGA 1577
Db 1574 AGCGGGGCGCTGTGGGCGCTGCTTCTGCTTACAGACATATCTTCCCTTCTCTTGA 1633
QY 1578 TTGCTATGATGTCAACAGAGAGAGCCAAATCGGAGACCCCGAGGCGACTGTATGGCA 1637
Db 1634 TTGCTATGATGTCAACAGAGAGAGCCAAATCGGAGACCCCGAGGCGACTGTATGGCA 1693
QY 1638 CATCTCCAGGTGAGCGAGGCTGCTGTGGCCCGGTAAGCCAGCATGCCATTCCTCG 1697
Db 1694 CATCTCCAGGTGAGCGAGGCTGCTGTGGCCCGGTAAGCCAGCATGCCATTCCTCG 1753
QY 1698 GGTATGCTGGGCGGCGCAGAGCTGGGCGCCAGGGGAAATTTGCTAAAGATCTTTCGGGCTG 1757
Db 1754 GGTATGCTGGGCGGCGCAGAGCTGGGCGCCAGGGGAAATTTGCTAAAGATCTTTCGGGCTG 1813
QY 1758 GGGATGTTTTCTTCAACACTGGGGAAGCTGTGCTGTGCGATGACCAAGTTTTCTCGCT 1817
Db 1814 GGGATGTTTTCTTCAACACTGGGGAAGCTGTGCTGTGCGATGACCAAGTTTTCTCGCT 1873
QY 1818 TCCATGATCTGATCTGGAACACCTTCAGGTGGAAGGGGGAAGATGTGGCCAAACCGAG 1877
Db 1874 TCCATGATCTGATCTGGAACACCTTCAGGTGGAAGGGGGAAGATGTGGCCAAACCGAG 1933
QY 1878 TGGCAGAGGCTTTCAGAGCCCTAGATTTCTTCAAGAGAGTGAACGTCTATGAGTCACTG 1937
Db 1934 TGGCAGAGGCTTTCAGAGCCCTAGATTTCTTCAAGAGAGTGAACGTCTATGAGTCACTG 1993
QY 1938 TGGCAGGCGATGAGGCGAGGCTGATGAGAGCCCTAGTTCTGTGCTCCCCCAAGCTT 1997
Db 1994 TGGCAGGCGATGAGGCGAGGCTGATGAGAGCCCTAGTTCTGTGCTCCCCCAAGCTT 2053
QY 1998 TGGACCTTATGAGCTTACACCCAGCTGTCTGAGAACTTTCGCACTTATCCCGGCGCC 2057
Db 2054 TGGACCTTATGAGCTTACACCCAGCTGTCTGAGAACTTTCGCACTTATCCCGGCGCC 2113
QY 2058 GATTCCTCAGGCTCAGAGAGTCTTTGGCCACACAGAGACCTTCAACAGCAGAAAGTTC 2117
Db 2114 GATTCCTCAGGCTCAGAGAGTCTTTGGCCACACAGAGACCTTCAACAGCAGAAAGTTC 2173
QY 2118 GGATGGCAATGAGGGCTTTCAGCCCGACAGCCGTCTGACCCACTGATACCTTTCGAGCC 2177
Db 2174 GGATGGCAATGAGGGCTTTCAGCCCGACAGCCGTCTGACCCACTGATACCTTTCGAGCC 2233
QY 2178 AGGCTGTAGGTCTCTACCTGCGCCCTCACAAGTGCAGGCGGTACAGGCGCTCTGCGAGAA 2237
Db 2234 AGGCTGTAGGTCTCTACCTGCGCCCTCACAAGTGCAGGCGGTACAGGCGCTCTGCGAGAA 2293

QY 2238 ACCTTCGATCTGAGAACTTCCACACTGAGAGCAGCTGAGAGAACTGTGGGTGG 2297
Db 2294 ACCTTCGATCTGAGAACTTCCACACTGAGAGCAGCTGAGAGAACTGTGGGTGG 2353
QY 2298 GGGCCGTTGAGGTGTACTGGGCTGTGAGGATCTTTCTATACCAAGACTGCGGTCACT 2357
Db 2354 GGGCCGTTGAGGTGTACTGGGCTGTGAGGATCTTTCTATACCAAGACTGCGGTCACT 2413
QY 2358 ATTTGTAATAATGAGGCTGAGCTGATCCAGCTGTCTGACCTA 2404
Db 2414 ATTTGTAATAATGAGGCTGAGCTGATCCAGCTGTCTGACCTA 2460
RESULT 15
ADA12301
ID ADA12301 standard; cDNA; 2574 BP.
XX
AC ADA12301;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human cDNA encoding secreted/transmembrane polypeptide PRO703.
XX
KW ss. gene; inflammatory disease; organ failure; atherosclerosis;
KW cardiac injury; infertility; birth defect; premature aging; AIDS; cancer;
KW diabetic complication; tissue typing; human.
XX
OS Homo sapiens.
XX
PN US2003055216-A1.
XX
PD 20-MAR-2003.
XX
PF 17-OCT-2001; 2001US-00978824.
XX
PR 21-MAY-1996; 96US-0018049P.
PR 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-00040220.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079656P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080333P.
PR 01-APR-1998; 98US-0080334P.
PR 01-APR-1998; 98US-0080337P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 08-APR-1998; 98US-0081195P.
PR 08-APR-1998; 98US-0081203P.
PR 09-APR-1998; 98US-0081229P.

```

PR 15-APR-1998; 98US-0081817P.
PR 15-APR-1998; 98US-0081819P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0081952P.
PR 15-APR-1998; 98US-0081955P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083332P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 29-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084598P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 13-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085573P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086023P.
PR 18-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 28-MAY-1998; 98US-0010541P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-0016897P.
PR 07-OCT-1998; 98WO-US02114P.
PR 06-NOV-1998; 98US-0018421P.
PR 06-NOV-1998; 98US-0018736P.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98WO-US02485P.
PR 07-DEC-1998; 98US-0020205P.
PR 22-DEC-1998; 98US-0021851P.
PR 22-DEC-1998; 98US-0113296P.
PR 23-DEC-1998; 98US-0113621P.
PR 05-JAN-1999; 99WO-US00010P.
PR 08-MAR-1999; 99US-0025446P.
PR 08-MAR-1999; 99WO-US00502P.
PR 10-MAR-1999; 99US-0026566P.

PR 10-MAR-1999; 99WO-US005190P.
PR 12-MAR-1999; 99US-0026721P.
PR 12-MAR-1999; 99US-0123957P.
PR 29-MAR-1999; 99US-0126773P.
PR 12-APR-1999; 99US-0028429P.
PR 12-APR-1999; 99US-0130232P.
PR 26-APR-1999; 99US-0131022P.
PR 28-APR-1999; 99US-0131455P.
PR 14-MAY-1999; 99US-0031832P.
PR 14-MAY-1999; 99US-0134287P.
PR 14-MAY-1999; 99WO-US01073P.
PR 02-JUN-1999; 99WO-US01225P.
PR 16-JUN-1999; 99US-0139557P.
PR 23-JUN-1999; 99US-0141037P.
PR 07-JUL-1999; 99US-0142680P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 25-AUG-1999; 99US-0038013P.
PR 25-AUG-1999; 99US-0038014P.
PR 25-AUG-1999; 99US-0162506P.
PR 29-OCT-1999; 99WO-US02831P.
PR 30-NOV-1999; 99WO-US02851P.
PR 02-DEC-1999; 99WO-US02855P.
PR 16-DEC-1999; 99WO-US03009P.
PR 30-DEC-1999; 99WO-US03124P.
PR 03-JAN-2000; 99WO-US00021P.
PR 06-JAN-2000; 99WO-US00027P.
PR 11-FEB-2000; 99WO-US00356P.
PR 18-FEB-2000; 99WO-US00434P.
PR 24-FEB-2000; 99WO-US00504P.
PR 02-MAR-2000; 99WO-US00584P.
PR 10-MAR-2000; 99WO-US00631P.
PR 21-MAR-2000; 99WO-US00753P.
PR 30-MAR-2000; 99WO-US00843P.
PR 17-MAY-2000; 99WO-US01370P.
PR 22-MAY-2000; 99WO-US01404P.
PR 30-MAY-2000; 99WO-US01494P.
PR 02-JUN-2000; 99WO-US01564P.
PR 28-JUL-2000; 99WO-US02071P.
PR 24-AUG-2000; 99WO-US02328P.
PR 08-NOV-2000; 99WO-US070923P.
PR 27-NOV-2000; 99WO-US072374P.
PR 01-DEC-2000; 99WO-US07367P.
PR 20-DEC-2000; 99WO-US07425P.
PR 20-DEC-2000; 99WO-US03495P.
PR 28-FEB-2001; 99WO-US00652P.
PR 22-MAR-2001; 99WO-US01674P.
PR 22-MAR-2001; 99WO-US01692P.
PR 22-MAR-2001; 99WO-US009552P.
PR 10-MAY-2001; 99WO-US085420P.
PR 10-MAY-2001; 99WO-US085428P.
PR 21-MAY-2001; 99WO-US017092P.
PR 01-JUN-2001; 99WO-US07203P.
PR 01-JUN-2001; 99WO-US01780P.
PR 05-JUN-2001; 99WO-US074503P.
PR 14-JUN-2001; 99WO-US08636P.
PR 19-JUN-2001; 99WO-US086342P.
PR 20-JUN-2001; 99WO-US019692P.
PR 29-JUN-2001; 99WO-US02106P.
PR 09-JUL-2001; 99WO-US02173P.
PR 30-JUL-2001; 99WO-US091858P.

XX (GETH ) GENENTECH INC.
XX PA
XX PI Ashkenazi AJ, Baker KP, Botstein D, Desnovers L, Baton DL;
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
Query Match 99.3%; Score 2387; DB 8; Length 2574;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 18 CTGCTCCGCCCGGAGTGGAGTGGAGGAGGCTGGGTGGGAATGGGCGGTGGCCAGCCGA 77
Db 74 CTGCTCTCGGCCCGGAGTGGAGTGGAGGAGGCTGGGTGGGAATGGGCGGTGGCCAGCCGA 133
QY 78 CGCGGCTCTCCGGAAGGAAGTCTCAGCTAGAAAGAGCGGCGCTTAGGTTTTCGGAAG 137
Db 134 CGCGGCTCTCCGGAAGGAAGTCTCAGCTAGAAAGAGCGGCGCTTAGGTTTTCGGAAG 193
QY 138 GAGGATTCAGGAGTGGTTTGGAGCGGCTGGAAACCAACAGGTGCCATTAAGAGAGGGGCT 197
Db 194 GAGGATTCAGGAGTGGTTTGGAGCGGCTGGAAACCAACAGGTGCCATTAAGAGAGGGGCT 253
QY 198 CCATGAGTGCCTCTCGTGTGCTGCCCTGTGCTGTGTCTACCGGCTGCTGTGTAAGC 257
Db 254 CCATGAGTGCCTCTCGTGTGCTGCCCTGTGCTGTGTCTACCGGCTGCTGTGTAAGC 313
QY 258 TACACTCTGAGCGGAGTGGGCTGGCTTCCGGCGGACTTGGGCTTTTGGGTCGAGCTC 317
Db 314 TACACTCTGAGCGGAGTGGGCTGGCTTCCGGCGGACTTGGGCTTTTGGGTCGAGCTC 373
QY 318 TGTGCTGCAAAAGGAGCTCTTCAGGCTCGGCGCTGCGCGGAGTGGCGCGGCGGAG 377
Db 374 TGTGCTGCAAAAGGAGCTCTTCAGGCTCGGCGCTGCGCGGAGTGGCGCGGCGGAG 433
QY 378 GTCCGAGGAGGAGCTGCAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 437
Db 434 GTCCGAGGAGGAGCTGCAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 493
QY 438 GCGACACCTTCTCTACAGGCTGCGGCGCTTACGTAAGTACTCAGAGGCGGAGGCGGAGA 497
Db 494 GCGACACCTTCTCTACAGGCTGCGGCGCTTACGTAAGTACTCAGAGGCGGAGGCGGAGA 553
QY 498 GTTACAGGAGCTGCAAGCGCTTCTTACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 557
Db 554 GTTACAGGAGCTGCAAGCGCTTCTTACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 613
QY 558 GCGACAGGAGGAGGAGGAGGAGCTGAGAGAGGCGAGCGGCGGAGGAGGAGGAGGAG 617
Db 614 GCGACAGGAGGAGGAGGAGGAGGAGCTGAGAGAGGCGAGCGGCGGAGGAGGAGGAGGAG 673
QY 618 CAGCGGCGGAGAGCGGCGGAGGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 677
Db 674 CAGCGGCGGAGAGCGGCGGAGGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 733
QY 678 CCGCGCGCTCTGTCTCAGCTGAGGAGCTGTGGGCTGTCTCTCCGCTGAGCCCAAGT 737
Db 734 CCGCGCGCTCTGTCTCAGCTGAGGAGCTGTGGGCTGTCTCTCCGCTGAGCCCAAGT 793
QY 738 TTCTGTGAGTGTGGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 797
Db 794 TTCTGTGAGTGTGGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 853
QY 798 CCTTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 857
Db 854 CCTTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 913
QY 858 TGGCGGCGGAGGAGTGTGAGTCCCTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 917
Db 914 TGGCGGCGGAGGAGTGTGAGTCCCTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 973
QY 918 TCCACTTGTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
Db 974 TCCACTTGTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1033
QY 978 AAGGTCCGCTGAATGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1037
Db 1034 AAGGTCCGCTGAATGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1093
QY 1038 CAGACAGGCTGCTGATCTTCACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1097
Db 1094 CAGACAGGCTGCTGATCTTCACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1153
QY 1098 TCAATCATGAGATCTGCAATGGCAGGAGCTTCTATCAGCTGTGAGGAGTCCACAGG 1157
Db 1154 TCAATCATGAGATCTGCAATGGCAGGAGCTTCTATCAGCTGTGAGGAGTCCACAGG 1213
QY 1158 AAGATGATCTACCTTGCGGCTCCCACTTACCAATGTCGAGTTCCTGCTGGGATCG 1217
Db 1214 AAGATGATCTACCTTGCGGCTCCCACTTACCAATGTCGAGTTCCTGCTGGGATCG 1273
QY 1218 TGGGCTGATGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1277
Db 1274 TGGGCTGATGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1333
QY 1278 TCTGGAGAGTTCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1337
Db 1334 TCTGGAGAGTTCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1393
QY 1338 GATACCTTGTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1397
Db 1394 GATACCTTGTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1453
QY 1398 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1457
Db 1454 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1513
QY 1458 AGGTGCTGAGAGATATGAGCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1517
Db 1514 AGGTGCTGAGAGATATGAGCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1573
QY 1518 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1577
Db 1574 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1633
QY 1578 TTGCTATGATATGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1637
Db 1634 TTGCTATGATATGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1693
QY 1638 CATCTCAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1697
Db 1694 CATCTCAGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1753
QY 1698 GATATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1757
Db 1754 GATATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1813
QY 1758 GGGATGTTTTCTTCAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1817
Db 1814 GGGATGTTTTCTTCAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1873
QY 1818 TCCATGATGCTATGAGAGACCTTCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1877
Db 1874 TCCATGATGCTATGAGAGACCTTCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1933
QY 1878 TGGCAGAGGCTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1937
Db 1934 TGGCAGAGGCTTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1993
QY 1938 TGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1997
Db 1994 TGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2053
QY 1998 TGGACCTTATGAGAGCTTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2057
Db 2054 TGGACCTTATGAGAGCTTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2113
QY 2058 GATTCTCAGGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2117
Db 2114 GATTCTCAGGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2173
QY 2118 GGTATGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2177
Db 2174 GGTATGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2233
QY 2178 AGGCTGTAGTGTCTACCTGCGGCTTCAACATGCTGCCGAGTACAGCGGCGCTCTGGAGAGGA 2237
```

Db	2234	AGGCTGTAGTGTGCTTACCTGCCCCCTCACACTGCCCCGGTACAGCGCCCTCCTGGCAGGAA	2293
Qy	2238	ACCTTCGAACTTGAGAACTTCCACACCTGAGGCACTGTGAGAGGAACTGTGTGGGGTGG	2237
Db	2294	ACCTTCGAATCTGAGAACTTCCACACCTGAGGCACTGTGAGAGGAACTGTGTGGGGTGG	2353
Qy	2298	GGGCGGTTGCAAGGTGACTGGGCTGTCAGGGAATCTTTTCTATACCAAACTGCGGTCACT	2357
Db	2354	GGGCGGTTGCAAGGTGACTGGGCTGTCAGGGAATCTTTTCTATACCAAACTGCGGTCACT	2413
Qy	2358	ATTTGTATATAATGTGGGTGAGCTGATCCAGCTGTCTCTGACCTA	2404
Db	2414	ATTTGTATATAATGTGGGTGAGCTGATCCAGCTGTCTCTGACCTA	2460

Search completed: July 2, 2004, 23:28:03
Job time : 954 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 17:32:25 ; Search time 21 Seconds
(without alignments)
3343.801 Million cell updates/sec

Title: US-10-030-226-2

Perfect score: 3843

Sequence: 1 MGVCQRTAPKPKESQLEKRA.....ALVPLTARYSALLAGLRI 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1257	32.7	689	1 JMW0107	very-long-chain ac
2	1056.5	27.5	646	1 AS5093	fatty acid transpo
3	1028	26.7	635	2 B87514	fatty acid transpo
4	903	23.5	608	2 H83284	probable very-long
5	858	22.3	655	2 H1879	hypothetical prote
6	841.5	21.9	597	2 D70609	probable fad6 pro
7	839.5	21.8	650	2 T21498	hypothetical prote
8	725.5	18.9	623	2 T43052	fatty acid transpo
9	654	17.0	622	2 S45899	probable membrane
10	405.5	10.6	522	2 H85484	probable crotonome
11	405.5	10.6	522	2 S40588	probable crotonome
12	393.5	10.2	522	2 S40588	probable carnitine
13	367	9.5	502	2 D70806	probable fad17 pr
14	353.5	9.2	517	2 A10510	probable crotonome
15	352	9.2	532	2 G70986	probable coa ligas
16	332	8.6	444	2 T50931	hypothetical prote
17	325	8.5	599	2 H72454	probable fatty-aci
18	313.5	8.2	632	2 B83163	probable AMP-bindi
19	311	8.1	584	2 C75364	probable long-chain
20	306	8.0	556	2 H81068	long-chain-fatty-a
21	302	7.9	506	2 G81798	long-chain-fatty-a
22	297.5	7.7	507	2 G70604	probable acid-CoA
23	295	7.6	564	2 C83327	probable AMP-bindi
24	293	7.5	508	2 H85271	probable long chain
25	291	7.6	541	2 T34850	probable acid-CoA
26	290.5	7.6	513	2 A69831	probable acid-CoA
27	286.5	7.5	515	2 F96011	probable long-chain
28	283	7.4	593	2 E69378	probable acid-CoA
29	278	7.2	1002	2 D87077	probable long-chain

30	276.5	7.2	562	2 AB0253	long-chain-fatty-a
31	273.5	7.1	561	2 B90743	long-chain-fatty-a
32	273.5	7.1	561	2 F85791	long-chain-fatty-a
33	273	7.1	524	2 H75530	probable acid-CoA
34	272.5	7.1	503	2 E70853	probable acid-CoA
35	269.5	7.0	561	1 S41589	long-chain-fatty-a
36	269.5	7.0	999	2 B70501	hypothetical prote
37	269	7.0	510	2 AB2782	long-chain fatty a
38	269	7.0	510	2 C97561	hypothetical prote
39	269	7.0	514	2 T46131	4-coumarate-CoA li
40	268	7.0	512	2 T41164	coenzyme a synthet
41	268	7.0	566	2 H82132	long-chain-fatty-a
42	267.5	7.0	639	2 AB3456	long-chain-fatty-a
43	265	6.9	562	2 B83233	long-chain-fatty-a
44	265.5	6.9	561	2 AC0725	long-chain-fatty-a
45	264	6.9	569	2 D82824	regulator of patho

ALIGNMENTS

RESULT 1

JMW0107 very-long-chain acyl-CoA synthetase related protein - mouse

N:Alternate names: VLCSR

C:Species: Mus musculus (house mouse)

C>Date: 10-Sep-1999 #sequence_rev: 10-Sep-1999 #text_change 08-Sep-2000

C/Accession: JMW0107

R/Berger, U.; Truppe, C.; Neumann, H.; Forst-Petter, S.

Biochem. Biophys. Res. Commun. 247, 255-260, 1998

A/Title: A novel relative of the very-long-chain acyl-CoA synthetase and fatty acid tran

A/Reference number: JMW0107; PMID:98308102; PMID:9642112

A/Accession: JMW0107

A/Molecule type: mRNA

A/Residues: 1689 <BBR>

A/Cross-references: GB:AJ223959

A/Experimental source: liver

C/Comment: This protein likely functions as a plasma membrane transporter of long chain

cids.

C:Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom

F;169-647/Domain: acetate-CoA ligase homology <ACL>

Query Match	32.7%	Score 1257;	DB 1;	Length 689;
Best Local Similarity	39.5%;	Pred. No. 3.8e-79;		
Matches 280;	Conservative 93;	Mismatches 257;	Indels 78;	Gaps 8;
QY	46	GSMAATLILPLLL-----LPLLLKIHMPQLRMLPADLAPVRA	86	
DB	37	GDPTCLVLLGLLGRPWISSMPPHMLSVGAALTLFLPLPPPLRLHNDVATFPKX	96	
QY	87	LCKRRALRAALAAADDEPGGCSLAWRLAELAQORAAHTFLHGSRRFSYSAERE	146	
DB	97	LFYGLKPRRR-----LNKHPPE-----TFV-----DALERQ	122	
QY	147	SNMAAARFLALGMDGPDG-----GDSGGSAGEGERAARGADAAAGSAGEPAGDGA	202	
DB	123	-----ALAW-----PDVALVCTGSESSSTTNSQDLARSQAAWVLAKKDKDAVIGN	169	
QY	203	RGGAAPLSPGATVALLPAGPEFLMLWFLAKAGIRTAFTALRRGLHLCRSCGA	262	
DB	170	TRMAALIVLPSKTIALS-----VFLGLAKGCPVAMINPHSGPFLHVSASSGA	221	
QY	263	RAIVLAPPELESLEPDLPALRAMGLHMAAGPETHAGISDLAEVSAEVDGVPYGLS	322	
DB	222	SVIVDPDQENIEVLPKLLAENIHCFYLGHSPPRGVEALGASJDAAPSDVPVPSLGA	281	
QY	323	PGSITDCLVYIFTSGTGLPKARISHLKIIOCGGFYOLCGVHQEDVITLALPLVMSSG	382	
DB	282	TIIMKSPATIFITSGTGLPKRALISHERVIVQSNVLSGCGADVVVDVPLVITIDL	341	
QY	383	LIGIVGMIIGATVILKSFSAQGFEDQGRVTVFYIIGLCRYLVNOPSKEGRHK	442	
DB	342	VLFGLGLQVATGATCVLAPKFSASRFWAECRQHGVTILVYGETLRYLCNVPEPDKIHT	401	

Db 547 LVTEGD--FDLEAFAYVDEALPSYARPLFVRLQKAIEFTTGFKYRKIDLVNNEGSDPSKT 604
QY 699 SPLLYLDQAVGAYLPLTTARYSALLAGNLR1 730
Db 605 KDPLFRDPKAG-YVKITKTKICAKIEGGGFRLL 635

RESULT 4

H83284
probable very-long-chain acyl-CoA synthetase PA2893 [imported] - Pseudomonas aeruginosa
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C/Accession: H83284
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: H83284
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-608 <STO>
A/Cross-references: GB:AE004715; GB:AE004091; NID:93948977; PIDN:AA06281.1; GSPDB:GN001
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA2893
C/Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom

Query Match 23.5%; Score 903; DB 2; Length 608;
Best Local Similarity 35.3%; Pred. No. 1e-54;
Matches 223; Conservative 80; Mismatches 243; Indels 86; Gaps 10;

QY 109 EGGCSLARLAEIAQORAAHTFLIGSRFSYSEARSRNAPAFRLALGMDWGDPOGD 168
Db 39 EKSLSLGALERRAALYDNPVLDGQRIRISALFNGMANRLARF----- 84
QY 169 GEGSAGEGERAPAGDAAAGSGAEFAGDGAARGGAAAPLPGATVALLPAGBEFL 228
Db 85 -----KAEGVGH-----GSVAVALLENREVELL 106
QY 229 WLMEGLAKGLRTAFTVPLRRGPLHCLRSOGARALVLAPELSELPDLPALRAGLH 288
Db 107 AILAAALAKGALGALVNTTQKGLVHSLNLYKGFHVGELREVEEYHVLGNAGH 166
QY 289 LM-----AAGPHTPAGISDL-----AEVSAEVDGVPVGYLSPOSTTDTCLYIFT 335
Db 167 CYWVDDGDTLGDPPGSPFGWMENLMRLAAGQTSNLEDTGRV-----RLKSCFYIYT 218
QY 336 SGTTLGLPKAARISHKILQCOGFR--QUCVHQBVDVYTLALPLYHMSGSLIGIVCGKIG 393
Db 219 SGTTLGLPKASIMSHKMTKAYGAGFGHSGLGGRDVLTLPLCYHNNAVVTCWSAALAGG 278
QY 394 ATVVLSKFSASGOEWEDCQHRVTYFOYIGELCRVLNQPSPKARSGHKKVALVAGSLRP 453
Db 279 AAMALRRFSSASGFVKDQHYRATCFYIGELCRVLNQPSCARERKSLTCTMGNGIRP 338
QY 454 DTWERFVRFPGLVLEYGLTEGNVATINTGORGAVGRASMLYKHIFPPLIRIVDTT 513
Db 339 SIMWFKRFRFRIORTTEFYASSEGNIGFTNPFNDVTGFSPPATY-----AIVRYDLEN 392
QY 514 GEPTRDPQGHCMATSPGEGLLVAPVSOQSPFLFYAGGPELAQGLLKDVRPQDVFNNT 573
Db 393 DRPRDAGFMFKYKGGVGLLISISKMPFDGT-DPAKSEAVILRDVFKGDAWNT 451
QY 574 GDLLVCDQGF--LRFHRTGDTFRWKGENVATTEVAEVEALDFLGEVNVYGVATVGEH 631
Db 452 GDLN--RDIGFKHQFVRLGDTFRWKGENVSTEVEALGAFDGVEDAVVYGVZIRPTN 509
QY 632 GRAGKAALVLRPPLALDMLQYTHVSENLPRYARPRFLRLOESLATTETFRQCYVRANE 691
Db 510 GRCKAALRLADGVLELSDALAAHLDRELPAVATPVFLRLREVEETGTGTYKKTDLKRD 569

QY 692 GFDPSTISDPLYLDQAVGAYLPLTTARYSAL-723
Db 570 AYDPARVSDKLFVRLPGSAGYQPLDAILYQAL 601

RESULT 5

h15879
hypothetical protein D1009.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C/Accession: T15879
R/Favell, T.
submitted to the EMBL Data Library, November 1995
A/Description: The sequence of C. elegans cosmid D1009.
A/Reference number: Z18422
A/Accession: T15879
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-655 <FAV>
A/Cross-references: EMBL:U40938; NID:g1072169; PID:g1072173; PIDN:AA81698.1; CESP:D1009
C/Genetics:
A/Gene: CESP:D1009.1
A/Intons: 2/2; 60/2; 125/3; 192/1; 344/2; 399/1; 547/3; 569/3; 602/1
C/Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom
F/128-611/Domain: acetate-CoA ligase homology <ACT>

Query Match 22.3%; Score 858; DB 2; Length 655;
Best Local Similarity 37.9%; Pred. No. 1.5e-51;
Matches 197; Conservative 95; Mismatches 200; Indels 28; Gaps 12;

QY 214 GATVALLPAPGPEFLWLMEGLAKGLRTAFTVPLRRGPLHCLRSOGARALV-----LAP 269
Db 132 GDVVALFMENSIDFFALWGLSKIGVSAFINSNLTLEPLASHINSKCKSCITNINLP 191
QY 270 EPLSELPDLPALRAGLHMAAG---PHTPAGISDLAEVSAEVDGVPVGYLSPOST 326
Db 192 MFXAREKKNLISDE---IHVFLAGTVDRHNSLQDL--HFSDEDPVIGL---NF 242
QY 327 TDTCLYIFTSGTGLPKAARISHKILQ--COGFRVQCGVHQBVDVYTLALPLYHMSGSLG 385
Db 243 RSVLCYIYISGTGNKRAVIMKFRFYAMAGKAFGINKSDVYITPMHSAAGING 302
QY 386 IYVCMGIGATVLSKFSASGOEWEDCQHRVTYFOYIGELCRVLNQPSPKARSGHKKVL 445
Db 303 IGSILAFGSTAVVRRKFSASNFKDCVKYNVATYVIGELCRVLNANPCPEKKQHVRLL 362
QY 446 AVSSGLRPDTWERFVRFPGLVLEYGLTEGNVATINTGORGAVGRASMLYKHIF--F 502
Db 363 MKNGLRGQIWKEFVGRFGIKIGLYSTEGNSIVNVDNHVAGCGMP--IYPHIGSLY 421
QY 503 PFLIRIVDTTGTGPIRDPQGHCMATSPGEGLLVAPVSOQSP--PFLFYAGGPELAQGL 559
Db 422 FVRLIKVDATAGLEBDKNGLCVPCVPGETGEVGVYKXKDLKKEGVSSGDTAK-KI 480
QY 560 LKDVRRPGVFNFTGGLVCDQGFPLRFHRTGDTFRWKGENVATTEVAEVEALDFLGE 619
Db 481 YRVFNGHGXVRSAGGILHMDLGYLYFVDRCGDFFRMKGENVSTTEVGGIIPQVNDVAD 540
QY 620 VNVYGVATVPEGHGRAMALVLRPPLALD--LMQYTHVSENLPRYARPRFLRLOESLAT 677
Db 541 ATYGVATVCKMGRAMAGIIVKDGVDVKFIADITSRLTENLASALIVFIRLCEVDR 600
QY 678 TETFKQCYVRANEGFD-BSTISDPLYLDQAVGAYLPLT 716
Db 601 TGTFFKLKXTDLQKGYDLVACKDPITYWSAAEKSYKPLT 640

RESULT 6

D70609
probable fad6 protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 18-Aug-2000
C/Accession: D70609

A:Molecule type: DNA
 A:Residues: 1-643 <OES>
 A:Cross-references: EMBL:Y15939; NID:G2687848; PTDN:CAA75802.1; PID:G2687850
 A:Experimental source: strain C2 (ATCC 48329)
 C:Genetics:
 A:Gene: fat1
 A:Introns: 169/3; 478/2
 C:Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom
 F:121-599/Domain: acetate-CoA ligase homology <AOL>

Query Match 18.9%; Score 725.5; DB 2; Length 643;
 Best Local Similarity 30.6%; Pred. No. 2,1e-42;
 Matches 194; Conservative 86; Mismatches 259; Indels 95; Gaps 12;

```

QY 120 ELAQAARATPLIHSSRRPSYSEARESENRAARALRLGMDWPGDGSSEGSAGGER 179
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 82 EOVQKQSNHPLIFEGKTKWSTYKESEAYTRVANNLIDPL----- 120
QY 180 AAFGAGDAAGSGAFAGDGAARCGGAAPLSPGATVALLPAGPEFLMWLPGIARAGL 239
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 -----DVQVGHMAVDGNSA----- 150
QY 240 RTAFVPTLRGRLHLCRSGARALVLPFLESLEPDLPALRAMGLHMAAPGTH-- 297
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 151 ATSFNNMNLTGAGLHCHKCECFVADIDIKANIEFCRGELEETGINT-----HY 203
QY 298 -PAGISDLAEVSAEVDGVPVGYLSSPOSITDTCLYFTSGTGLPKARISHLKIQCQ 356
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 204 DPFSTSSL-----PNNPPIPSRRENIELDSVGLIYTSGTGLPKGVFISTGRLE 257
QY 357 -GYQLCGVHQBVDIYLLPLPHNSGSLIGIVCMGIGATVYLKSKPSAGQFMEDCQ 415
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 258 WSISKYKMKPDRMYTCMPLYHAAHSLCTASVHGGGIVVLSNKSFKKFEVAVASE 317
QY 416 VTVPQYIGELCRYVNOPPSKAEGRKYLAVGSLRDPDTERFVRPRGPIQLVET 475
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 318 ANITVYSELGRYLLNGPKSPYDRAHKQVMKGMGRDVEARERENITIIHELPAAT 377
QY 476 EG-----NVATINTGORGAVGRASWLYKHIFPSSLIRYVTTGEPRI 526
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 378 DGLGMRNRNAGPFPAFCIALRGLT-----MWMKFRNGEVLVKMDLDEIWRDNGFAIR 433
QY 527 TSPGPRGLVAVPVGOSPFGLVAGGP-----LAQGLLDVPRPGVFNPTGDLV 580
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 434 CAVNPEGQMLPRLPETH---LAGAPSYNNETATQSRITDVEQKGLMKSSDMKRD 489
QY 581 DQGLRPFHRTGDTFRMKGENVATTEVAEVEALDLEOVNVVGYTVPGHGRAGMALV 640
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 490 AEGRYVFDRLGDTFRMKSENVSTNEVADVGTFRQIAETVYGVLPVNGNGRVSINCH 549
QY 641 LRPPHALV--MQLYTHVSENLPPYARPRFRLQESLATTTTFKQCKVRANEGP 698
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 550 GRRRDVRVIDIRPAALAKHARDRLPGYAVPLFLRVTPALEYTGTLKIQKRLKQ 609
QY 699 S--DPLVYLDQAVGAYLPLTTARYSALLAGNLRI 730
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 610 SGEDELVYLPFGSDIYLPFGKMEWGIVDKRIRL 643

```

RESULT 9
 S45899
 Probable membrane protein YBR041W - Yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein YBR0411
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 02-Sep-2000
 C:Accession: S45899
 R:Andr , B.; Czapluch, C.; Hein, C.; Jauniaux, J.C.; Urrestarazu, A.; Vissers, S.
 submitted to the Protein Sequence Database, August 1994
 A:Reference number: S45893
 A:Accession: S45899
 A:Molecule type: DNA
 A:Residues: 1-643 <AND>
 A:Cross-references: EMBL:Z35910; NID:G536263; PIDN:CAA84983.1; PID:G536264; MIPS:YBR041W

A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:FAT1
 A:Cross-references: SGD:S0000245; MIPS:YBR041W
 A:Map position: 2R
 C:Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom
 C:Keywords: transmembrane protein
 F:1-27/Domain: transmembrane #status predicted <TM1>
 F:55-71/Domain: transmembrane #status predicted <TM2>
 F:133-623/Domain: acetate-CoA ligase homology <ACU>
 F:149-167/Domain: transmembrane #status predicted <TM3>
 F:304-322/Domain: transmembrane #status predicted <TM4>

Query Match 17.0%; Score 654; DB 2; Length 623;
 Best Local Similarity 33.7%; Pred. No. 1,8e-37;
 Matches 169; Conservative 84; Mismatches 193; Indels 56; Gaps 14;

```

QY 214 GATVALLPAGPEFLMWFLGAKAGLRTAFVPTALRGLHLCRSGARALVLP-- 269
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 137 GDVVAIDCTNKPFLFVFLMWLMWIGAIAPFLNNTKGTFLVSLKISNTVOVFIDPAGN 196
QY 270 ---EFLSELPDLRALRANGHLHMAAPGTHPAGISDLAEVSAEVDGVPVGYLS----- 321
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 197 PIRESEBETNALPDVKNLYLE-----EODLMHEL---LNSQSPFLQDNVR 241
QY 322 SPOSITD--TCLYFTSGTGLPKARISHLK-ILOCGFYOICGVHQBVDIYLLPLVH 378
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 242 TPLGLTDFKRSMLIYSGTGLPKSAIMGRSSVGCQFEGVLMHTNISTVFTAMPPLH 301
QY 379 MGSLLIGVCMGIGATVYLKSKPSAGQFMEDCQHRVTVFQYIGELCRYVNOPPSKAE 438
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 302 STALLGACALISHGGLALSHKFSASTFWKQYLLGAHIDYGVGCAYLHTPISKYE 361
QY 439 RGHKYLAVGSLRDPDTERFVRPRGPIQLVETYGGLTEGNVATINTGORGAVGRAS--- 495
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 362 KMKKIVAVANGRLPRIMQDFRKRFEIVIGERYATENARPATTE--QKDPGIGACGN 419
QY 496 ----WLYKHIFPSSLIRYVDTTGEPI-RDPQGHMATSPEGGLVLA---PVSGQSP 544
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 420 YGTIIQWFLS--FOQTLVMDPNDSDVIYRNSKGEVAVPGEPGMLRKIRFPKKPETS 477
QY 545 FLGYPAGPELAGQKLKDVPRGDFVFNPTGDLVCCDQGLFLPHDPTGTPFRKGENVAT 604
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 478 FQGITLNAKETSKVADVFRGCDANYRCGDLKADYGLMTFTLDMGDTFRKKSENVST 537
QY 605 TEVAEVEALDPLQ--EVNVVGYTVPGHGRAGMALVLRPPHALDMLQLYTHVSE--- 658
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 538 TEVEDQLTSSNKQVAGVAVGIVKVPYEGRAGFAVYKL-TNSLDITAKTKLINDSLR 596
QY 659 -NPPYARPRFRLQESLATTE 679
   ||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 597 LNPSTYAMPLEVKFVDEIKMTD 618

```

RESULT 10
 H90633
 Probable crotonobetaine/carnitine-CoA ligase [imported] - *Escherichia coli* (strain O157:1)
 C:Species: *Escherichia coli*
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: H90633
 R:Hayashi, T.; Makino, K.; Ohishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
 gasawara, N.; Yasunaga, T.; Kohara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-23, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genom
 A:Reference number: A99629; M0ID:21156251; PMID:11283796
 A:Accession: H90633
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-522 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA833463.1; PID:G13359496; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: Ecs0040

C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 10.6%; Score 405.5; DB 2; Length 522;
Best Local Similarity 25.8%; Pred. No. 2,2e-20;
Matches 154; Conservative 78; Mismatches 235; Indels 131; Gaps 19;

```

QY 110 GGCSLAWRLAEIAQQRAHTPLIHGS-----RRFSYSEARESNFAAFLRALGMDWGP 164
DB 10 GGQHLRQMDLADYVGHKTLICSSGGVNRVSYLEINQINRTANLFF-TLG----- 63
QY 165 DGDSDSGSAGEGERAAPGADAAAGSGAEFAGGGAARGGGAAPLSPGATVALLPAG 224
DB 64 -----IRKGDVALLHLDNC 77
QY 225 PEFLIMFGLAKGLTAFTAVPTALRRGPLLHCLRSCGAPALVLA---PEFLSELEPDL 280
DB 78 PEFLFCWFGTAKIGAIWVFNARLLREESAMILQNSQACLVTSQAFYMYQOIQOEDAT 137
QY 281 ALRAGMLHMAAGPGTHPAGISDLAEVSAEVDGPVGY-----LSSPQSTID 328
DB 138 QLR---HI-----CLTDVALPADDGVSFTQKNOQPATLCYAPPLSTD 179
QY 329 TCLYITSGTGLPKAARISHLKILOCQFYQL--CGVQEDVYIALPLYMSSGLIGI 386
DB 180 TAEILFTSGTTRPKGVVITHYN-LRFAGYSAQCALRDVDVYLVTPAFHIDQOCTAA 238
QY 387 VCGMGIGATVVLKSKFSAGQFEDCCQHRVTYFOYIGELCRYLVNQPSPKARGHKVR-L 445
DB 239 MAFSAGATFVLEKYSAPAFMGQVQKXATTECIPMMIRLTMQPSANDRQRLREV 298
QY 446 AVSGSLRPDTERFVRFPGLQVLETYGLTEGNVATINTYQOR-----GAVGRASWLY 498
DB 299 MYLINSEQKDTFCERFG-VRLTISYGMTETIVGII--GDRPGDKRWPISIGRAGFCY 354
QY 499 KHIFPPLIRYDVTGEPTRDPQGHCMATSPGPG-LTVAAPSQSPFLGYAGPBLAOG 557
DB 355 -----DAEIRD--DNRRLPAGEIGEICIKGYPKGTIFREYFLNKR-ATA 396
QY 558 KLLKDVFRPDVFNFTGDLVLCDDQGFLEFHDRTGDFPMKGENVATTEVAEVEALDPL 617
DB 397 KYLE-----ADGWLHTGDTGYRDEGFFYFIDRCNMIRKGENVSCVLENIITATHPKI 451
QY 618 QEVNYYGVTVPGHSGAGMAALVLRPPHALDMQLYTHSENLPYARPRRLLOESL 675
DB 452 QDIVVVGIR-DSIRDEAIKAFVVLNEGTELSSEBFFRCQNMAKFKVSYLEIRKDL 508

RESULT 11
H85484
probable crotonobetaine/carnitine-CoA ligase caic [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C/Accession: H85484
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A/Reference number: A85480; MUID:21074935; PMID:11206551
A/Accession: H85484
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <STO>
A/Cross-references: GB:AEO05174; NID:g12512724; PIDN:AAG54340.1; GSPDB:GN00145; UMGF:Z0C
A/Experimental source: strain O157:H7, substrain EDJ933
C:Genetics:
A:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 10.6%; Score 405.5; DB 2; Length 522;
Best Local Similarity 25.8%; Pred. No. 2,2e-20;
Matches 154; Conservative 78; Mismatches 235; Indels 131; Gaps 19;
110 GGCSLAWRLAEIAQQRAHTPLIHGS-----RRFSYSEARESNFAAFLRALGMDWGP 164

```

```

DB 10 GGQHLRQMDLADYVGHKTLICSSGGVNRVSYLEINQINRTANLFF-TLG----- 63
QY 165 DGDSDSGSAGEGERAAPGADAAAGSGAEFAGGGAARGGGAAPLSPGATVALLPAG 224
DB 64 -----IRKGDVALLHLDNC 77
QY 225 PEFLIMFGLAKGLTAFTAVPTALRRGPLLHCLRSCGAPALVLA---PEFLSELEPDL 280
DB 78 PEFLFCWFGTAKIGAIWVFNARLLREESAMILQNSQACLVTSQAFYMYQOIQOEDAT 137
QY 281 ALRAGMLHMAAGPGTHPAGISDLAEVSAEVDGPVGY-----LSSPQSTID 328
DB 138 QLR---HI-----CLTDVALPADDGVSFTQKNOQPATLCYAPPLSTD 179
QY 329 TCLYITSGTGLPKAARISHLKILOCQFYQL--CGVQEDVYIALPLYMSSGLIGI 386
DB 180 TAEILFTSGTTRPKGVVITHYN-LRFAGYSAQCALRDVDVYLVTPAFHIDQOCTAA 238
QY 387 VCGMGIGATVVLKSKFSAGQFEDCCQHRVTYFOYIGELCRYLVNQPSPKARGHKVR-L 445
DB 239 MAFSAGATFVLEKYSAPAFMGQVQKXATTECIPMMIRLTMQPSANDRQRLREV 298
QY 446 AVSGSLRPDTERFVRFPGLQVLETYGLTEGNVATINTYQOR-----GAVGRASWLY 498
DB 299 MYLINSEQKDTFCERFG-VRLTISYGMTETIVGII--GDRPGDKRWPISIGRAGFCY 354
QY 499 KHIFPPLIRYDVTGEPTRDPQGHCMATSPGPG-LTVAAPSQSPFLGYAGPBLAOG 557
DB 355 -----DAEIRD--DNRRLPAGEIGEICIKGYPKGTIFREYFLNKR-ATA 396
QY 558 KLLKDVFRPDVFNFTGDLVLCDDQGFLEFHDRTGDFPMKGENVATTEVAEVEALDPL 617
DB 397 KYLE-----ADGWLHTGDTGYRDEGFFYFIDRCNMIRKGENVSCVLENIITATHPKI 451
QY 618 QEVNYYGVTVPGHSGAGMAALVLRPPHALDMQLYTHSENLPYARPRRLLOESL 675
DB 452 QDIVVVGIR-DSIRDEAIKAFVVLNEGTELSSEBFFRCQNMAKFKVSYLEIRKDL 508

RESULT 12
S40558
probable carnitine-CoA ligase (EC 6.2.1.-) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Mar-2002
C/Accession: E64724; S40558; I41013
R:Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: E64724
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-522 <BLAT>
A/Cross-references: GB:AEO00114; GB:U00096; NID:g1786217; PIDN:AACT3148.1; PID:g1786221;
A/Experimental source: strain K-12, substrain MG1655
R:Yura, T.; Mortl, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobu
Mol. Microbiol. 13, 775-786, 1994
A>Title: Molecular characterization of the cat operon necessary for carnitine metabolism
A/Reference number: I41010; MUID:95115548; PMID:7815937
A/Accession: I41013
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA

```

A:Residues: 1-102, 'R', 104-256, 'PR', 259-290, 'R', 292-332, 'S', 334-349, 'A', 351-378, 'V', 380-3
 A:Cross-references: EMBL:X73904; NID:9563860; PDB:CAA5113.1; PID:9563864
 A:Experimental source: strain 044 K74
 C:Genetics:
 A:Gene: calc
 C:Function:
 A:Pathway: lysine degradation; tryptophan metabolism
 C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology
 C:Keywords: acid-thiol ligase; ATP; phosphoprotein; transmembrane protein
 F:63-521/Domain: acetate-CoA ligase homology <ACI>
 F:81-97/Domain: transmembrane #status predicted <TM1>
 F:123-251/Domain: transmembrane #status predicted <TM2>
 F:194/Active site: Lys #status predicted

Query Match 10.2% Score 393.5; DB 2; Length 522;
 Best Local Similarity 25.7%; Pred. No. 1.5e-19;
 Matches 155; Conservative 79; Mismatches 227; Indels 143; Gaps 22;

QY 110 GGCSIAWLAELIAQRAHFLIHG-----RRFSYSAERESNAPAPFLALGMDWG 164
 DB 10 GGOHLRQWMDLADYGHKHTALICSSGVNRYSYLHNGEINTNLFY-TLG----- 63
 QY 165 DGGSDGEGSAGSGEAAAGADAAAGSAGFRAGGGAARGGAAPLSPGATVALLPAG 224
 DB 64 -----IKGDKVALHLDNC 77
 QY 225 PEFLWLMFGLAKAGLRTAFAVPTALRRGPLHC-----LRSCGARALVLA---PEFLS 274
 DB 78 PEFICWGLAKIG--ALMVPIAR-----LICEESAMILQNSQCLLYTSQPFEMVQI 131
 QY 275 LEPDLPARAMGLHMAAGPGTHPAGISDLAEVSAEVDGPYGY-----LSS 322
 DB 132 QOEDATQLR---HI-----CLTDVALPADDDGVSSPTQKNQOPATLCAAP 173
 QY 323 POSIDTCLYFTSGTGLPKARISHLKIIQCCGFGYL--CGVQGEVDVIALPIVMS 380
 DB 174 PLSTDDTAEILFTSGTSPKGVVITHN-LRFAQYSAMQALRDDVYITVPAFHID 232
 QY 381 GSLLGIVGCMGIGATVILKSKFSAGQFWDCCQHRVTVOYIGLCRYLVNQPSPKARG 440
 DB 223 CQCTAAMAFAFSAGATFVVEKYSARAFWQYQKYRATVTECIPIIMRTLMQPSANDQ 292
 QY 441 HKVR-LAVSGGLRPPTMERFVRPGLQVETVGLTEGNVATINYGR-----GANG 492
 DB 293 HLRVVMYTLNSEDKDAFCERFG-VRLTISYGTETVGI---GDRPDKRWBSIG 348
 QY 493 RASWLKAIFFPSLIRYDVTGTEPRIDPOGHCMATSPGPG-LIAPVSOQSPFLGYAG 551
 DB 349 RVGFCE-----ABIRD--DNRRLPAGEIGEICIKGIPGTFKRYFLN 391
 QY 552 PELAGSKLKXVFRGDFVFNIGDLVCCDQGLRFHRTGDTFRMKENATIEVAVF 611
 DB 392 FO-ATAKYLE-----ADGMLHGTGYPDEDFEYVDRRCNMKRGGENNSCVELNII 445
 QY 612 EALDFLQEVNRYGVTPVPEHGRAGMAALVLRPPHALDMQVTVHSENLPRYAPRFLRL 671
 DB 446 AAHPKIDIVVVGIR-DSIRDEAIGAFVYLVNGETLSEEFRCCEQMMAFKPSTYIEI 504
 QY 672 QESL 675
 DB 505 RKDL 508

RESULT 13
 D70806
 Probable faadD7 protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 18-Aug-2000
 C:Accession: D70806
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:9825987; PMID:9634230
 A:Accession: D70806
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-502 <COL>
 A:Cross-references: GB:AL022022; GB:AL123456; NID:93261554; PDB:CAA1743.1; PID:g292444
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: faadD7
 C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology
 F:51-486/Domain: acetate-CoA ligase homology <ACL>

Query Match 9.5% Score 367; DB 2; Length 502;
 Best Local Similarity 28.4%; Pred. No. 9.8e-18;
 Matches 151; Conservative 66; Mismatches 214; Indels 98; Gaps 22;

QY 200 GAARGGAAAPLSPGAT--VALLPAGPEFLWLMFGLAKAGRTAVPTAL---RRGPL 254
 DB 39 GAALAAALRRLDPPRPHGVLLQNTPF---SATVAGALSGVPGLVNVRG--- 91
 QY 255 HCLRSCARALVLAPEFLSELEPDLALRAMGLHMAAGPGTHPAGISDL-----LAE 307
 DB 92 -----MALAGDI-AKADQQLVLGSSGSAEVDVDEHINVDSPMTDE 132
 QY 308 VSAEVDGPVGYLSPQSIDTCLYFTSGTGLPKARISHLKI-LQCCGFGYQLCGVHQ 366
 DB 133 VAHRTDEV-RFSA--DLADLFMLFTSGTSDPFAVKCSHRKVAIAGTTIQRSIGR 189
 QY 367 EDVITYALPIYHNSGILG--IVGCMGIGATVILKSKFSAGQFWDCCQHRVTVOYIG 423
 DB 190 DVCYVSMPLFHSNVALVGVAAACQG---SWALRRKFSASQFLADVRRYGATVANYG 246
 QY 424 ELGRVYNQPPSAEGHKKRLAVGSLGDPDTERFVRPGLQVETVGLTEGNVA--TI 482
 DB 247 KPLSYVLATPELDDNDNPRAYGNBVGPDIDRRRGCV-VMDGSGSTGGVATIR 305
 QY 483 NYTGORGAVRASWLKHIFFPSLIRYDVTGTEPRIDPOGHCMATSPG-----PGLV 536
 DB 306 TLDTPPAGALP-----LPGSIQVDDTGP-----CPTGVGELVNTAGSG--- 347
 QY 537 APVSQSPFLGYAG--GPELAGSKLKXVFRGDFVFNIGDLVCCDQGLRFHRTGDT 594
 DB 348 -----GREGYNDEAAEAE-----RANGVYHSGDLAYRDAGVAFAGRLGWM 391
 QY 595 FRWKGENVATTEVAEVEALDFLQEVNRYGVTPVPEHGRAGMAALVLRPPHALDMQVLT 654
 DB 392 MRVDGNLGTAPRIERLVMPDTEVAVYFVPDP-VYGQVMAALVLAQTGDAKFKRA 450
 QY 655 HVSE--NLPRYAPRFLRLOESLATTETFKQCKRYANSGPDBSTISDPLY 703
 DB 451 FLTEQPDLGHKQWPSYRVASGLPRMTFKVIRQLSAGEV--ACADPVM 498

RESULT 14
 A10510
 Probable crotonobetaine/carnitine-CoA ligase [imported] - Salmonella enterica subsp. ent.
 C:Species: Salmonella enterica subsp. enterica serovar Typh
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: A10510
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, C.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: A10510
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-517 <PAR>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 17:23:58 ; Search time 60 Seconds
(without alignments)
3437.661 Million cell updates/sec

Title: US-10-030-226-2

Perfect score: 3843

Sequence: 1 MGVCORTRAPWKEKSQLERA.....AYLPITARYSALLAGNLR I 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseqp29Jan04:*
1: geneseqp1990s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3843	100.0	730	2	AAV41699
2	3843	100.0	730	3	AAAB44255
3	3843	100.0	730	3	AAAB24054
4	3843	100.0	730	4	AAAB60388
5	3843	100.0	730	4	AAAB83779
6	3843	100.0	730	6	AAAB25201
7	3843	100.0	730	6	AAAB72207
8	3843	100.0	730	6	AAAB84115
9	3843	100.0	730	6	AAAB84887
10	3843	100.0	730	6	AAAB61085
11	3843	100.0	730	6	AAAB80354
12	3843	100.0	730	6	AAAB24641
13	3843	100.0	730	6	AAAB19656
14	3843	100.0	730	6	AAAB12302
15	3843	100.0	730	6	AAAB19547
16	3843	100.0	730	6	AAAB73608
17	3843	100.0	730	7	AAAB76324
18	3843	100.0	730	7	AAAB43750
19	3843	100.0	730	7	AAAB61510
20	3843	100.0	730	7	AAAB63474
21	3843	100.0	730	7	AAAB66574
22	3843	100.0	730	7	AAAB68698
23	3843	100.0	730	7	AAAB62758
24	3843	100.0	730	7	AAAB67823
25	3843	100.0	730	7	AAAB41143

26	3843	100.0	730	7	ADCB7198	ADCB7198 Human sec
27	3843	100.0	730	7	ADCB2134	ADCB2134 Human sec
28	3843	100.0	730	7	ADCB1767	ADCB1767 Human sec
29	3843	100.0	730	7	ADCB9136	ADCB9136 Human sec
30	3843	100.0	730	7	ADCB5190	ADCB5190 Human sec
31	3843	100.0	730	7	ADCB16304	ADCB16304 Human sec
32	3843	100.0	730	7	ADCB2919	ADCB2919 Human sec
33	3843	100.0	730	7	ADCB72277	ADCB72277 Human sec
34	3843	100.0	730	7	ADCB16928	ADCB16928 Human sec
35	3843	100.0	730	7	ADCB48436	ADCB48436 Human sec
36	3843	100.0	730	8	ADCB9537	ADCB9537 Human sec
37	3843	100.0	811	6	AAAB32069	AAAB32069 Human TRI
38	3683	95.8	700	4	AAAB95135	AAAB95135 Human pro
39	3657	95.2	702	2	AAAB14969	AAAB14969 Amino ac1
40	3657	95.2	702	4	AAAB83278	AAAB83278 Human FAT
41	3592	93.5	683	4	AAAB62172	AAAB62172 Human gen
42	3525.5	91.7	759	6	AAAB32083	AAAB32083 Human TRI
43	3474	90.4	667	5	AAAB63572	AAAB63572 Human alD
44	2759	71.8	613	2	AAAB14933	AAAB14933 Amino ac1
45	2759	71.8	613	4	AAAB83271	AAAB83271 Murine FA

ALIGNMENTS

RESULT 1	AAV41699	AAV41699 standard; protein; 730 AA.
ID	AAV41699	
XX	AAV41699;	
AC	AAV41699;	
XX	07-DEC-1999 (first entry)	
DT		
XX		
DE	Human PRO703 protein sequence.	
XX		
KW	Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;	
KM	probe; blood coagulation disorder; cancer; cellular adhesion disorder;	
XX	secreted protein; transmembrane protein.	
OS	Homo sapiens.	
XX		
PN	WO9946281-A2.	
XX		
PD	16-SEP-1999.	
XX		
PF	08-MAR-1999;	99WO-US005028.
XX		
PR	10-MAR-1998;	98US-0077450P.
PR	11-MAR-1998;	98US-0077632P.
PR	11-MAR-1998;	98US-0077641P.
PR	11-MAR-1998;	98US-0077649P.
PR	12-MAR-1998;	98US-0077791P.
PR	13-MAR-1998;	98US-0078004P.
PR	17-MAR-1998;	98US-00040220.
PR	17-MAR-1998;	98US-0078886P.
PR	20-MAR-1998;	98US-0078910P.
PR	20-MAR-1998;	98US-0078936P.
PR	20-MAR-1998;	98US-0078939P.
PR	25-MAR-1998;	98US-0079294P.
PR	26-MAR-1998;	98US-0079656P.
PR	27-MAR-1998;	98US-0079663P.
PR	27-MAR-1998;	98US-0079664P.
PR	27-MAR-1998;	98US-0079689P.
PR	27-MAR-1998;	98US-0079728P.
PR	27-MAR-1998;	98US-0079786P.
PR	30-MAR-1998;	98US-0079920P.
PR	30-MAR-1998;	98US-0079923P.
PR	31-MAR-1998;	98US-0080105P.
PR	31-MAR-1998;	98US-0080107P.
PR	31-MAR-1998;	98US-0080165P.
PR	31-MAR-1998;	98US-0080194P.
PR	01-APR-1998;	98US-0080327P.
PR	01-APR-1998;	98US-0080328P.

PR 01-APR-1998; 98US-0080333P.
 PR 01-APR-1998; 98US-0080334P.
 PR 08-APR-1998; 98US-0081049P.
 PR 08-APR-1998; 98US-0081070P.
 PR 08-APR-1998; 98US-0081071P.
 PR 09-APR-1998; 98US-0081195P.
 PR 09-APR-1998; 98US-0081203P.
 PR 09-APR-1998; 98US-0081229P.
 PR 15-APR-1998; 98US-0081817P.
 PR 15-APR-1998; 98US-0081838P.
 PR 15-APR-1998; 98US-0081952P.
 PR 15-APR-1998; 98US-0081955P.
 PR 21-APR-1998; 98US-0082568P.
 PR 21-APR-1998; 98US-0082569P.
 PR 22-APR-1998; 98US-0082700P.
 PR 22-APR-1998; 98US-0082704P.
 PR 23-APR-1998; 98US-0082804P.
 PR 23-APR-1998; 98US-0082867P.
 PR 23-APR-1998; 98US-0082966P.
 PR 27-APR-1998; 98US-0083336P.
 PR 28-APR-1998; 98US-0083322P.
 PR 29-APR-1998; 98US-0083392P.
 PR 29-APR-1998; 98US-0083495P.
 PR 29-APR-1998; 98US-0083496P.
 PR 29-APR-1998; 98US-0083499P.
 PR 29-APR-1998; 98US-0083500P.
 PR 29-APR-1998; 98US-0083545P.
 PR 29-APR-1998; 98US-0083546P.
 PR 29-APR-1998; 98US-0083588P.
 PR 29-APR-1998; 98US-0083589P.
 PR 30-APR-1998; 98US-0083742P.
 PR 05-MAY-1998; 98US-0084366P.
 PR 06-MAY-1998; 98US-0084414P.
 PR 06-MAY-1998; 98US-0084415P.
 PR 07-MAY-1998; 98US-0084598P.
 PR 07-MAY-1998; 98US-0084600P.
 PR 07-MAY-1998; 98US-0084627P.
 PR 07-MAY-1998; 98US-0084637P.
 PR 07-MAY-1998; 98US-0084639P.
 PR 07-MAY-1998; 98US-0084640P.
 PR 07-MAY-1998; 98US-0084643P.
 PR 13-MAY-1998; 98US-0085323P.
 PR 13-MAY-1998; 98US-0085388P.
 PR 13-MAY-1998; 98US-0085399P.
 PR 15-MAY-1998; 98US-0085573P.
 PR 15-MAY-1998; 98US-0085579P.
 PR 15-MAY-1998; 98US-0085800P.
 PR 15-MAY-1998; 98US-0085822P.
 PR 15-MAY-1998; 98US-0085892P.
 PR 15-MAY-1998; 98US-0085972P.
 PR 15-MAY-1998; 98US-0085700P.
 PR 15-MAY-1998; 98US-0085702P.
 PR 18-MAY-1998; 98US-0086023P.
 PR 22-MAY-1998; 98US-0086322P.
 PR 22-MAY-1998; 98US-0086323P.
 PR 22-MAY-1998; 98US-0086414P.
 PR 22-MAY-1998; 98US-0086430P.
 PR 22-MAY-1998; 98US-0086466P.
 PR 28-MAY-1998; 98US-0087098P.
 PR 28-MAY-1998; 98US-0087106P.
 PR 28-MAY-1998; 98US-0087208P.
 PR 30-UTL-1998; 98US-0094651P.
 PR 11-SEP-1998; 98US-0100038P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
 XX
 DR WFI, 1999-551358/46.
 DR N-PSDB; AA23977.
 XX
 PT New secreted and transmembrane polypeptides and their polymuclectides,
 PT useful for treating blood coagulation disorders, cancers and cellular
 PT adhesion disorders.

XX
 PS Claim 12; Fig 39; 530pp; English.
 XX
 CC The present invention describes secreted and transmembrane polypeptides
 CC and their polymuclectides. The nucleotide sequences are useful as sources
 CC of probes, primers, for chromosome mapping, and for generation of
 CC antisense sequences. They can also be used to create transgenic animals.
 CC The proteins can be used to treat a variety of diseases and disorders,
 CC depending on their function. Diseases that may be treated include blood
 CC coagulation disorders, cancers and cellular adhesion disorders. They may
 CC also be used to raise antibodies. AA23891 to AA23438, and AA41681 to
 CC AA41774 represent polynucleotide and polypeptide sequence given in the
 CC exemplification of the present invention
 XX
 SQ Sequence 730 AA;
 Query Match 100.0%; Score 3843; DB 2; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGVCQTRAPWKEKSLERAPALGFRGSGMFGSGMGTVPVEEGSMALLPLPLLL 60
 DB 1 MGVCQTRAPWKEKSLERAPALGFRGSGMFGSGMGTVPVEEGSMALLPLPLLL 60
 QY PLLLLKLHLPOLRWLPADLAPAVRALCCRALRARALAAAAADPEPGGCSLAWRLAE 120
 DB PLLLLKLHLPOLRWLPADLAPAVRALCCRALRARALAAAAADPEPGGCSLAWRLAE 120
 QY 61 PLLLLKLHLPOLRWLPADLAPAVRALCCRALRARALAAAAADPEPGGCSLAWRLAE 120
 DB 61 PLLLLKLHLPOLRWLPADLAPAVRALCCRALRARALAAAAADPEPGGCSLAWRLAE 120
 QY 121 LAQORAAHTFLIGSRSPFSYSEKERSNRAARFLPALMDMGPDGDSGSGAGGGERA 180
 DB 121 LAQORAAHTFLIGSRSPFSYSEKERSNRAARFLPALMDMGPDGDSGSGAGGGERA 180
 QY 181 APGAGDAAAGSGAEFAGDGAAGCGAAAPLSGATVALLPAGPEFLIMFGLAKAGUR 240
 DB 181 APGAGDAAAGSGAEFAGDGAAGCGAAAPLSGATVALLPAGPEFLIMFGLAKAGUR 240
 QY 241 TAFVPTALRRGPLLHCLIRSCGARALVLABPESLEPDLPALRANGLHMAAGPGRHPAG 300
 DB 241 TAFVPTALRRGPLLHCLIRSCGARALVLABPESLEPDLPALRANGLHMAAGPGRHPAG 300
 QY 301 ISDLLAEVSAEVGVPVGVYSSPQSTTPTCLVYFTSGTGLPKAARISHLKILCOGFFYQ 360
 DB 301 ISDLLAEVSAEVGVPVGVYSSPQSTTPTCLVYFTSGTGLPKAARISHLKILCOGFFYQ 360
 QY 361 LCGVHOEDVLYALPYHMSGSLIGVCMGIGATVYLKSKFSAGQFWEDCOOHRVTYQ 420
 DB 361 LCGVHOEDVLYALPYHMSGSLIGVCMGIGATVYLKSKFSAGQFWEDCOOHRVTYQ 420
 QY 421 YIGELCRYLYNQPSSAERGHKRLAVSGLRPDIWERFRKRGPIQVLETYGLTGNVA 480
 DB 421 YIGELCRYLYNQPSSAERGHKRLAVSGLRPDIWERFRKRGPIQVLETYGLTGNVA 480
 QY 481 TINYTGORGAVGRASWLYKAIFFPSLIRYDVTGGEIRDPQHCWATSPEGELVAPVS 540
 DB 481 TINYTGORGAVGRASWLYKAIFFPSLIRYDVTGGEIRDPQHCWATSPEGELVAPVS 540
 QY 541 QOSPFPLGAGGPELAGCKLKDVFRRGPDVFRNTGDLVCDQGFRLFPHRTGDPFKWKE 600
 DB 541 QOSPFPLGAGGPELAGCKLKDVFRRGPDVFRNTGDLVCDQGFRLFPHRTGDPFKWKE 600
 QY 601 NVATTEVAEVEALDFEOEVNNGVTVPGHEGAGAAALVREPHALDLMQVTHYSENI 660
 DB 601 NVATTEVAEVEALDFEOEVNNGVTVPGHEGAGAAALVREPHALDLMQVTHYSENI 660
 QY 661 PPYARPPFLRLQSLATTEFFKQOKVMANEGDPSLTSLDPLVLDQAVGAYPLITARY 720
 DB 661 PPYARPPFLRLQSLATTEFFKQOKVMANEGDPSLTSLDPLVLDQAVGAYPLITARY 720
 QY 721 SALLAGNLR 730
 DB 721 SALLAGNLR 730

RESULT 2
 AAB44255 standard; protein; 730 AA.
 AC AAB44255;
 DT 08-FEB-2001 (first entry)
 DE Human PRO703 (UNQ367) protein sequence SEQ ID NO:102.
 KW Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
 KW expressed sequence tag; detection; cancer.
 OS Homo sapiens.
 XX MO200053756-A2.
 XX WO200053756-A2.
 PD 14-SEP-2000.
 XX 18-FEB-2000; 2000WO-US004341.
 XX 08-MAR-1999; 99WO-US005028.
 XX 12-MAR-1999; 99US-0123957P.
 XX 29-MAR-1999; 99US-0126773P.
 XX 21-APR-1999; 99US-0130233P.
 XX 28-APR-1999; 99US-0131445P.
 XX 14-MAY-1999; 99US-0134287P.
 XX 23-JUN-1999; 99US-0141037P.
 XX 26-JUL-1999; 99US-0145698P.
 XX 29-OCT-1999; 99US-0162506P.
 XX 30-NOV-1999; 99WO-US028313.
 XX 02-DEC-1999; 99WO-US028551.
 XX 16-DEC-1999; 99WO-US030035.
 XX 30-DEC-1999; 99WO-US031243.
 XX 30-DEC-1999; 99WO-US031274.
 XX 05-JAN-2000; 2000WO-US000219.
 XX 06-JAN-2000; 2000WO-US000277.
 XX 06-JAN-2000; 2000WO-US000376.
 XX (GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Bolstein D, Desnoyers L, Eaton DL;
 XX Ferrare N, Filvaroff E, Fong S, Gao W, Geber H, Gerritsen ME;
 XX Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan MJ;
 XX K11avin JF, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;
 XX Stewart TM, Tumas D, Williams PM, Wood WI;
 XX WPI; 2000-611443/58.
 XX N-PSDB; AACT8481.
 XX Novel PRO polypeptides and polynucleotides used in detection methods, to
 XX target bioactive molecules to specific cells, and to modulate cellular
 XX activities.
 XX Claim 12, Fig 39, 636p; English.
 XX AACT8458 to AACT8599 represent polynucleotide and EST (expressed sequence
 XX tag) sequences which encode secreted or transmembrane PRO polypeptides.
 XX The PRO polynucleotides and polypeptides have cytosolic activity. The
 XX polynucleotides and polypeptides can be used for detecting the presence
 XX of PRO polypeptides in samples, for linking bioactive molecules to cells
 XX and for modulating biological activities of cells, using the polypeptides
 XX for specific targeting. The polypeptide targeting can be used to kill the
 XX target cells, e.g. for the treatment of cancers. The polypeptide pairs
 XX provide specific targeting of bioactive molecules to cells. AACT8600 to
 XX AACT8887 represent PCR primers and probes used in the isolation of the
 XX PRO polynucleotide sequences
 XX Sequence 730 AA;
 Query Match 100.0%; Score 3843; DB 3; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGVCORTBAPMKESQLEERALLGFRKGGSGMPSGNNQVTEPEAGSMALLPLPLLL 60
 Db 1 MGVCORTBAPMKESQLEERALLGFRKGGSGMPSGNNQVTEPEAGSMALLPLPLLL 60
 QY 61 PLILLKILHMPQLRWLPADLAFARALCKRRALRABALAAAAADPEGCGSLAWRLAE 120
 Db 61 PLILLKILHMPQLRWLPADLAFARALCKRRALRABALAAAAADPEGCGSLAWRLAE 120
 QY 121 LAQORAAHFTLHGSRRFSYSEAEBSNRARAFALMDMDPDGDSGSGAGGGERA 180
 Db 121 LAQORAAHFTLHGSRRFSYSEAEBSNRARAFALMDMDPDGDSGSGAGGGERA 180
 QY 181 APGAGDAAAGSGAEFAGDGAARGGGAAPLSFGATVALLPAGPEFLMWGLAKAGLR 240
 Db 181 APGAGDAAAGSGAEFAGDGAARGGGAAPLSFGATVALLPAGPEFLMWGLAKAGLR 240
 QY 241 TAFVPTALRRGPLHCLRSAGARALVLAPEFLSEFPDLPALRAMLHMAAGPGTHPAG 300
 Db 241 TAFVPTALRRGPLHCLRSAGARALVLAPEFLSEFPDLPALRAMLHMAAGPGTHPAG 300
 QY 301 ISDLAEVSAEVDGPVPGYLTSSPOSTITDCLYIFTSGTGLPKAAISHLKIQCQGFYQ 360
 Db 301 ISDLAEVSAEVDGPVPGYLTSSPOSTITDCLYIFTSGTGLPKAAISHLKIQCQGFYQ 360
 QY 361 LCGVHOEDVIYIALPLYHNSGSLGIVGCMGIGATVILKSFSAQGFMEDCCQHRVTFQ 420
 Db 361 LCGVHOEDVIYIALPLYHNSGSLGIVGCMGIGATVILKSFSAQGFMEDCCQHRVTFQ 420
 QY 421 YIGELCRYVNOPPSAERGHKRYLAVSGSLPDMTERVRRFPGPOVLETYGLGNGVA 480
 Db 421 YIGELCRYVNOPPSAERGHKRYLAVSGSLPDMTERVRRFPGPOVLETYGLGNGVA 480
 QY 481 TIINYTGQRGAVGRASNLTKHIFPFSILIRYVTTGSEIRDPQCHCATSPGEGELLVAPVS 540
 Db 481 TIINYTGQRGAVGRASNLTKHIFPFSILIRYVTTGSEIRDPQCHCATSPGEGELLVAPVS 540
 QY 541 QOSPFIVGAGPEPLAAGKLLKDVFRPGDVFFNTGDLVDDGFLRFHRTGDTFRWKKE 600
 Db 541 QOSPFIVGAGPEPLAAGKLLKDVFRPGDVFFNTGDLVDDGFLRFHRTGDTFRWKKE 600
 QY 601 NVATTEVAEVEFALDPLQEVNTYGVTPGHEGRAGAAALVLRPPHALDLMQLYTHVSENL 660
 Db 601 NVATTEVAEVEFALDPLQEVNTYGVTPGHEGRAGAAALVLRPPHALDLMQLYTHVSENL 660
 QY 661 PPAARPRFLRLOESLATTETTFKQKVRANEGPDSTISDPLVLDQAVGAYLPLTTARY 720
 Db 661 PPAARPRFLRLOESLATTETTFKQKVRANEGPDSTISDPLVLDQAVGAYLPLTTARY 720
 QY 721 SALLAGNLRI 730
 Db 721 SALLAGNLRI 730
 RESULT 3
 AAB24054
 ID AAB24054 standard; protein; 730 AA.
 AC AAB24054;
 DT 25-JAN-2001 (first entry)
 DE Human PRO703 protein sequence SEQ ID NO:29.
 KW Human; tumour; diagnosis; neoplastic disease; identification; cancer;
 KW tumour; diagnosis; detection; neoplastic cell growth; proliferation;
 KW cytosolic; antiinflammatory; immunomodulatory; inflammatory disorder;
 KW immunological disorder.
 OS Homo sapiens.
 XX MO200053754-A1.

CC The invention relates to a novel human fatty acid transporter, PSEC67
 CC (AAB60388), and to cDNA encoding it (AA27417). PSEC67 is responsible for
 CC the uptake of oleic acid into cells. The invention also relates to
 CC vectors and host cells comprising a PSEC67 nucleic acid; the recombinant
 CC production of PSEC67; an antibody against PSEC67; methods of screening
 CC for compounds which can regulate the uptake of long-chain fatty acids
 CC into cells; and the compounds thus identified. The PSEC67 protein and the
 CC gene encoding it are useful as targets for the treatment or prevention of
 CC diseases associated with defective metabolism of long-chain fatty acids.
 CC Such diseases include cardiomyopathy, skeletal muscle disorders, or renal
 CC failure. PSEC67 nucleic acids may also be used in gene therapy for such
 CC disorders. The present sequence represents PSEC67

XX
 CC Sequence 730 AA;

Query Match 100.0%; Score 3843; DB 4; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTRAPWKEKSOLEERALGFRKSGSGMFASGNNQTVPIEAGSMALLLLPLILL 60
 DB 1 MGVCORTRAPWKEKSOLEERALGFRKSGSGMFASGNNQTVPIEAGSMALLLLPLILL 60
 QY 61 PLILIKLHLMFQRLWLPADLAFAYRALCCRRALRAALAAAAADPEGPGGCSLAWRLAE 120
 DB 61 PLILIKLHLMFQRLWLPADLAFAYRALCCRRALRAALAAAAADPEGPGGCSLAWRLAE 120
 QY 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFALALGMDWGPDDGSGSGAGEERA 180
 DB 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFALALGMDWGPDDGSGSGAGEERA 180
 QY 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFALALGMDWGPDDGSGSGAGEERA 180
 DB 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFALALGMDWGPDDGSGSGAGEERA 180
 QY 181 APGAGDAAGSAGAEFAGDGAARGGGAAAPLSPGATVALLPAPPEFLMWFGIAKAGLR 240
 DB 181 APGAGDAAGSAGAEFAGDGAARGGGAAAPLSPGATVALLPAPPEFLMWFGIAKAGLR 240
 QY 241 TAFVPTALRRGPPLLHCLRSAGARALVLAPEFLSLEPDLPALRMGLHLMWAGFTHPAG 300
 DB 241 TAFVPTALRRGPPLLHCLRSAGARALVLAPEFLSLEPDLPALRMGLHLMWAGFTHPAG 300
 QY 301 ISDLIAEVSADVDPGVTSSPQSIITDTCYIFPSGTGPKRARIISHKTLCCQGFYQ 360
 DB 301 ISDLIAEVSADVDPGVTSSPQSIITDTCYIFPSGTGPKRARIISHKTLCCQGFYQ 360
 QY 361 LCGVHOEDVIYLAFLYHMSGSLIGIVCGMIGATVVLKSKFSAGQFWEDECQHRVTVFQ 420
 DB 361 LCGVHOEDVIYLAFLYHMSGSLIGIVCGMIGATVVLKSKFSAGQFWEDECQHRVTVFQ 420
 QY 421 YIGELCRILVNOPPKSARGHKVRLAVSGSGIRPDTWEFVARRFPLQVLETVGLTEGNA 480
 DB 421 YIGELCRILVNOPPKSARGHKVRLAVSGSGIRPDTWEFVARRFPLQVLETVGLTEGNA 480
 QY 481 TINYTGQGAAGVARSMLYKHIFPESLIYDVDTGEPIRDPOGHCAVSPGEGLLVAPVS 540
 DB 481 TINYTGQGAAGVARSMLYKHIFPESLIYDVDTGEPIRDPOGHCAVSPGEGLLVAPVS 540
 QY 541 QOSPFLYAGGBELAQGLKLDVFRPGDFVFNFTGDLVCDQGLRFHRTGDTFRMKGE 600
 DB 541 QOSPFLYAGGBELAQGLKLDVFRPGDFVFNFTGDLVCDQGLRFHRTGDTFRMKGE 600
 QY 601 NVATEVEAVEFALDFLOAVNVYGVTVPGHGRAGMALVLRPPALDLMQLYTHVSNL 660
 DB 601 NVATEVEAVEFALDFLOAVNVYGVTVPGHGRAGMALVLRPPALDLMQLYTHVSNL 660
 QY 661 PEYARPRFLQESLATTETFEKQKVRMANEGFDSTLSDPLVYLDQAVGAYLPITARY 720
 DB 661 PEYARPRFLQESLATTETFEKQKVRMANEGFDSTLSDPLVYLDQAVGAYLPITARY 720
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730

RESULT 5

AA83279
 ID AAB83279 standard; protein; 730 AA.
 XX
 AC AAB83279;
 XX
 DT 06-JUL-2001 (first entry)
 XX
 DE Human FATP3 SEQ ID NO: 117.
 XX
 XX Fatty acid transport protein; FATP; human; mouse; rat; rice blast fungus;
 KM yeast; fat absorption; obesity; diabetes; heart disease; hyperlipidemia;
 KM weight control; tuberculosis; TB; anti-fungal.
 XX
 OS Homo sapiens.
 XX
 PN WO200121795-A2.
 XX
 PD 29-MAR-2001.
 XX
 PF 21-SEP-2000; 2000MO-US025891.
 XX
 PR 23-SEP-1999; 99US-00405504.
 PR 23-SEP-1999; 99US-00405505.
 PR 16-DEC-1999; 99US-00465280.
 PR 17-FEB-2000; 2000US-00506252.
 PR 06-JUL-2000; 2000US-00611197.
 XX
 PA (MHED) WHITEHEAD INST BIOMEDICAL RES.
 PA (MILL) MILLIENIUM PHARM INC.
 XX
 PI Stahl A, Hirsch DJ, Lodish HF, Gimeno RE, Tartaglia LA;
 DR WPI; 2001-354783/37.
 DR N-PSDB; AAF89054.
 XX
 PT New fatty acid transport proteins (FATPs) useful for the manufacture of
 PT medicament for treating obesity, diabetes and heart disease.
 PS
 PS Claim 31; Fig 112; 287pp; English.
 XX
 CC The present invention provides the protein and coding sequences of fatty
 CC acid transport proteins (FATPs) from a number of species, including
 CC FATP1, FATP2, FATP3, FATP4, FATP5 and FATP6 from the human, FATP1-FATP5
 CC from the mouse, FATP4 and b from C. elegans, and FATP from Aspergillus
 CC nidulans, Drosophila, zebrafish, Magnaporthe grisea, Mycobacterium
 CC tuberculosis and Cochliobolus heterostrophus. The FATP from M.
 CC tuberculosis can be used to identify inhibitors which can then be used to
 CC treat TB. That from M. grisea (also known as rice blast fungus) can be
 CC used to develop anti-fungal agents capable of preventing infection of
 CC rice. Those from the human can be used to develop treatments for
 CC diabetes, heart disease, obesity, hyperlipidemia and weight control. The
 CC present sequence is one of the sequences described in the exemplification
 CC of the invention
 XX
 SQ Sequence 730 AA;
 Query Match 100.0%; Score 3843; DB 4; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTRAPWKEKSOLEERALGFRKSGSGMFASGNNQTVPIEAGSMALLLLPLILL 60
 DB 1 MGVCORTRAPWKEKSOLEERALGFRKSGSGMFASGNNQTVPIEAGSMALLLLPLILL 60
 QY 61 PLILIKLHLMFQRLWLPADLAFAYRALCCRRALRAALAAAAADPEGPGGCSLAWRLAE 120
 DB 61 PLILIKLHLMFQRLWLPADLAFAYRALCCRRALRAALAAAAADPEGPGGCSLAWRLAE 120
 QY 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFALALGMDWGPDDGSGSGAGEERA 180
 DB 121 LAOQRAAHTFLIHGSRFRSYSEARESNRAARAFALALGMDWGPDDGSGSGAGEERA 180
 QY 181 APGAGDAAGSAGAEFAGDGAARGGGAAAPLSPGATVALLPAPPEFLMWFGIAKAGLR 240

DB 181 AAGAGAAAGSGAERFAGGGAARGGAAAPLSFGATVALLPAGPEFLMLWGLAKAELR 240
QY 241 TAFVPTALRGPGLHLCLRSRGARALVLAPEFLESLEPDLPALRAMGLHMAAGPCTHPAG 300
DB 244 TAFVPTALRGPGLHLCLRSRGARALVLAPEFLESLEPDLPALRAMGLHMAAGPCTHPAG 300
QY 301 ISDLAEVSAEVDGPPGGLSSPQSTITDTCLYIFTSSTTGPKAKRISHKLILQCGFYQ 360
DB 301 ISDLAEVSAEVDGPPGGLSSPQSTITDTCLYIFTSSTTGPKAKRISHKLILQCGFYQ 360
QY 361 LCGVHOEDVIYALPLPYHMSGSLIGVGMGIGATVLSKFSASAGQFEDCQHRVTYFQ 420
DB 361 LCGVHOEDVIYALPLPYHMSGSLIGVGMGIGATVLSKFSASAGQFEDCQHRVTYFQ 420
QY 421 YIGELCRILVNOPPSKAEKRGKRLAVSGSLRPDWEFRFRFGPLQVLELYGLTEGNA 480
DB 421 YIGELCRILVNOPPSKAEKRGKRLAVSGSLRPDWEFRFRFGPLQVLELYGLTEGNA 480
QY 481 TINYTGORGAVRASMLYKHPFSLIRDYDTGEPIDPOGHCAATSPGEPGLLVAVS 540
DB 481 TINYTGORGAVRASMLYKHPFSLIRDYDTGEPIDPOGHCAATSPGEPGLLVAVS 540
QY 541 QOSPFLGYAGGPELAQGLKLDVFRPGDVFNNTGDLVCDQGFIRFHDRTGDTFRWKE 600
DB 541 QOSPFLGYAGGPELAQGLKLDVFRPGDVFNNTGDLVCDQGFIRFHDRTGDTFRWKE 600
QY 601 NVATTEVAVFEPFALPFLQEVNYYGVTPGHGSRAGMALVLRPPHALDMLQTYHVSNTL 660
DB 601 NVATTEVAVFEPFALPFLQEVNYYGVTPGHGSRAGMALVLRPPHALDMLQTYHVSNTL 660
QY 661 PPYAPRFLRLQESLATTETFKQOKVRMANEGFDBSTLSDPLVYLDAVGYLDTTARY 720
DB 661 PPYAPRFLRLQESLATTETFKQOKVRMANEGFDBSTLSDPLVYLDAVGYLDTTARY 720
QY 721 SALLAGNLR 730
DB 721 SALLAGNLR 730

RESULT 6
ABO25201
ID ABO25201 standard; protein; 730 AA.
XX ABO25201;
AC ABO25201;
XX 09-SEP-2003 (first entry)
DE Novel human secreted and transmembrane protein PRO703.
XX
XX Human, secreted and transmembrane protein; PRO; virucide; gene therapy;
KW cell death; growth induction cascade; blood coagulation cascade;
XX
XX Homo sapiens.
OS
XX
XX US2003050239-A1.
PN
XX 13-MAR-2003.
PD
XX
PF 15-OCT-2001; 2001US-00978191.
XX
XX 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064248P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 11-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-00040220.

PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079656P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080333P.
PR 01-APR-1998; 98US-0080334P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 09-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081203P.
PR 09-APR-1998; 98US-0081229P.
PR 15-APR-1998; 98US-0081817P.
PR 15-APR-1998; 98US-0081819P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0081952P.
PR 15-APR-1998; 98US-0081955P.
PR 21-APR-1998; 98US-0082688P.
PR 21-APR-1998; 98US-0082689P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083336P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083546P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084441P.
PR 07-MAY-1998; 98US-0084589P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 13-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085573P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085689P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086023P.

PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086466P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091010P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98US-002021141.
PR 02-NOV-1998; 98US-00184216.
PR 06-NOV-1998; 98US-00187368.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98US-00202054.
PR 07-DEC-1998; 98US-00218517.
PR 22-DEC-1998; 98US-0113296P.
PR 23-DEC-1998; 98US-0113621P.
PR 05-JAN-1999; 99US-00254465.
PR 08-MAR-1999; 99US-00254465.
PR 10-MAR-1999; 99US-00265686.
PR 10-MAR-1999; 99US-00265686.
PR 12-MAR-1999; 99US-00267213.
PR 12-MAR-1999; 99US-0123957P.
PR 29-MAR-1999; 99US-0126773P.
PR 12-APR-1999; 99US-00284291.
PR 21-APR-1999; 99US-0130232P.
PR 26-APR-1999; 99US-0131022P.
PR 28-APR-1999; 99US-0131445P.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99US-0134287P.
PR 14-MAY-1999; 99US-0010773.
PR 02-JUN-1999; 99US-0012252.
PR 16-JUN-1999; 99US-0139557P.
PR 23-JUN-1999; 99US-0141037P.
PR 07-JUL-1999; 99US-0142680P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380142.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99US-00282313.
PR 02-DEC-1999; 99US-0028551.
PR 02-DEC-1999; 99US-0028551.
PR 16-DEC-1999; 99US-00301243.
PR 30-DEC-1999; 99US-00301274.
PR 03-DEC-1999; 99US-00301274.
PR 06-JAN-2000; 2000US-0000219.
PR 06-JAN-2000; 2000US-0000219.
PR 11-FEB-2000; 2000US-0000376.
PR 18-FEB-2000; 2000US-0000341.
PR 24-FEB-2000; 2000US-0000504.
PR 02-MAR-2000; 2000US-00005841.
PR 10-MAR-2000; 2000US-00006319.
PR 21-MAR-2000; 2000US-00007532.
PR 30-MAR-2000; 2000US-00008439.
PR 17-MAY-2000; 2000US-00013705.
PR 22-MAY-2000; 2000US-00014042.
PR 30-MAY-2000; 2000US-00014941.
PR 02-JUN-2000; 2000US-00015264.
PR 28-JUN-2000; 2000US-00015264.
PR 24-AUG-2000; 2000US-00020710.
PR 08-NOV-2000; 2000US-00023328.
PR 27-NOV-2000; 2000US-00023749.
PR 01-DEC-2000; 2000US-00023749.

PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000US-00747259.
PR 28-FEB-2001; 2001US-0006520.
PR 28-MAR-2001; 2001US-00816744.
PR 22-MAR-2001; 2001US-00816920.
PR 22-MAR-2001; 2001US-00816920.
PR 22-MAR-2001; 2001US-00816920.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001US-0017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001US-00872035.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001US-0019692.
PR 29-JUN-2001; 2001US-0021066.
PR 09-JUL-2001; 2001US-0021735.
PR 30-JUL-2001; 2001US-00918585.
XX
XX (GETH) GENENTECH INC.
PI Ashkenazi AJ, Baker KP, Botstein D, Denoyers J, Eaton DJ,
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
Query Match 100.0%; Score 3843; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVCQRTAPWKEKSLERBALGFRKSGSGMFGAGNQTPIEBAGSMALLPLLL 60
DQ 1 MGVCQRTAPWKEKSLERBALGFRKSGSGMFGAGNQTPIEBAGSMALLPLLL 60
QY 61 PLILLKLIHMPQIRWLPADLAFVRLCCRRALRAALAAAADPBGSGSLARLAE 120
DQ 61 PLILLKLIHMPQIRWLPADLAFVRLCCRRALRAALAAAADPBGSGSLARLAE 120
QY 121 LAQRAAHTFLIHGSRFYSSEARESNRABAFRLALGMDGPDGDSGSGAGERA 180
DQ 121 LAQRAAHTFLIHGSRFYSSEARESNRABAFRLALGMDGPDGDSGSGAGERA 180
QY 181 APAGDPAAGSGAEFGGGAARGGGAAPLSBQATVALLLPAGPFLMLWGLARAR 240
DQ 181 APAGDPAAGSGAEFGGGAARGGGAAPLSBQATVALLLPAGPFLMLWGLARAR 240
QY 241 TAFVPTALRRGPIHLCLRSCGARALVLAPEFLESLEPDPALRAMGLHMAAGPGR 300
DQ 241 TAFVPTALRRGPIHLCLRSCGARALVLAPEFLESLEPDPALRAMGLHMAAGPGR 300
QY 301 ISDLAEVAEAVDGPVGYLSBPQITDTCLYIFTSIGTGLPKAARISHKLILQCGFYQ 360
DQ 301 ISDLAEVAEAVDGPVGYLSBPQITDTCLYIFTSIGTGLPKAARISHKLILQCGFYQ 360
QY 361 LCGVHOEDVIYALPLTHWSSGLGVGCMGIGATVLLSKSASAGFMWDCQHRVTYQ 420
DQ 361 LCGVHOEDVIYALPLTHWSSGLGVGCMGIGATVLLSKSASAGFMWDCQHRVTYQ 420
QY 421 YIGELCRYLVNQPSPKAEGRGHKRLAVGSLRPDTEREVRBFGPQVLETYGLTEGNA 480
DQ 421 YIGELCRYLVNQPSPKAEGRGHKRLAVGSLRPDTEREVRBFGPQVLETYGLTEGNA 480
QY 481 TIYITGQRAVGRASVLYKHIFPFSILRDVYTTGEBIRDPQCHMAATSGEGGLVAP 540
DQ 481 TIYITGQRAVGRASVLYKHIFPFSILRDVYTTGEBIRDPQCHMAATSGEGGLVAP 540
QY 541 QOSPFLGAGPFLAAGKLIKDFRPGDVFFNTGDLVDCDDGFLRFHRTDITFRMKE 600
DQ 541 QOSPFLGAGPFLAAGKLIKDFRPGDVFFNTGDLVDCDDGFLRFHRTDITFRMKE 600
QY 601 NVATTEVAEVEFALDPLQEVNNTYGVTPGHGGRAGMAALVLRPHALDLMQLYTHSEN 660
DQ 601 NVATTEVAEVEFALDPLQEVNNTYGVTPGHGGRAGMAALVLRPHALDLMQLYTHSEN 660
QY 661 PFAARPRFLRLOESLATTETFKQOKRVMANEGPDSTLSDPYLVLDQAVGAVLPLTARY 720
DQ 661 PFAARPRFLRLOESLATTETFKQOKRVMANEGPDSTLSDPYLVLDQAVGAVLPLTARY 720

DB 661 PYAAPRRRLQESLTTTETFKQKVRMANESGFDSTLSDPLVLDQVAGVLPITARY 720
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730

RESULT 7
 ID ABU72207 standard; protein; 730 AA.
 XX ABU72207;
 AC
 XX
 DT 16-JUN-2003 (first entry)
 DE Novel human secreted and transmembrane protein PRO703.
 XX
 KW Human; secreted and transmembrane protein; PRO; antiinflammatory;
 KW antiarteriosclerotic; cardiatic; anti-infertility; anti-HIV; cytosclastic;
 KW antidiabetic; gene therapy; inflammatory disease; organ failure;
 KW atherosclerosis; cardiac injury; infertility; birth defect;
 KW premature aging; AIDS; cancer; diabetic complication; chromosome mapping;
 KW gene mapping; pharmaceutical; diagnostic; biosensor; bioreactor;
 KW tissue typing.
 XX
 OS Homo sapiens.
 XX
 PN US2002192706-A1.
 XX
 PD 19-DEC-2002.
 XX
 PF 24-OCT-2001; 2001US-00999832.
 XX
 PR 17-OCT-1997; 97US-0062250F.
 PR 13-NOV-1997; 97US-0064249F.
 PR 13-NOV-1997; 97US-0065311P.
 PR 21-NOV-1997; 97US-0066364F.
 PR 10-MAR-1998; 98US-0077450P.
 PR 11-MAR-1998; 98US-0077632P.
 PR 11-MAR-1998; 98US-0077641P.
 PR 11-MAR-1998; 98US-0077649P.
 PR 12-MAR-1998; 98US-0077791P.
 PR 13-MAR-1998; 98US-0078004P.
 PR 17-MAR-1998; 98US-0004022O.
 PR 20-MAR-1998; 98US-0078886P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079294P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079689P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 27-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 31-MAR-1998; 98US-0080105P.
 PR 31-MAR-1998; 98US-0080107P.
 PR 31-MAR-1998; 98US-0080165P.
 PR 31-MAR-1998; 98US-0080194P.
 PR 01-APR-1998; 98US-0080327P.
 PR 01-APR-1998; 98US-0080328P.
 PR 01-APR-1998; 98US-0080333P.
 PR 01-APR-1998; 98US-0080334P.
 PR 08-APR-1998; 98US-0081049P.
 PR 08-APR-1998; 98US-0081070P.
 PR 08-APR-1998; 98US-0081071P.
 PR 09-APR-1998; 98US-0081159P.
 PR 09-APR-1998; 98US-0081203P.
 PR 09-APR-1998; 98US-0081229P.
 PR 15-APR-1998; 98US-0081817P.

PR 15-APR-1998; 98US-0081819P.
 PR 15-APR-1998; 98US-0081838P.
 PR 15-APR-1998; 98US-0081952P.
 PR 15-APR-1998; 98US-0081955P.
 PR 21-APR-1998; 98US-0082568P.
 PR 21-APR-1998; 98US-0082569P.
 PR 22-APR-1998; 98US-0082700P.
 PR 22-APR-1998; 98US-0082704P.
 PR 22-APR-1998; 98US-0082797P.
 PR 22-APR-1998; 98US-0082804P.
 PR 23-APR-1998; 98US-0082796P.
 PR 07-OCT-1998; 98WO-US021141.
 PR 20-NOV-1998; 98WO-US024855.
 PR 08-MAR-1999; 99WO-US000106.
 PR 10-MAR-1999; 99WO-US000528.
 PR 14-MAY-1999; 99WO-US010733.
 PR 02-JUN-1999; 99WO-US012252.
 PR 30-NOV-1999; 99WO-US028313.
 PR 02-DEC-1999; 99WO-US028551.
 PR 02-DEC-1999; 99WO-US028565.
 PR 16-DEC-1999; 99WO-US030095.
 PR 30-DEC-1999; 99WO-US031243.
 PR 30-DEC-1999; 99WO-US031274.
 PR 05-JAN-2000; 2000WO-US000219.
 PR 06-JAN-2000; 2000WO-US000277.
 PR 06-JAN-2000; 2000WO-US000376.
 PR 11-FEB-2000; 2000WO-US003565.
 PR 18-FEB-2000; 2000WO-US004341.
 PR 24-FEB-2000; 2000WO-US005004.
 PR 02-MAR-2000; 2000WO-US005841.
 PR 10-MAR-2000; 2000WO-US006219.
 PR 21-MAR-2000; 2000WO-US007539.
 PR 30-MAR-2000; 2000WO-US008439.
 PR 17-MAY-2000; 2000WO-US013705.
 PR 22-MAY-2000; 2000WO-US014042.
 PR 30-MAY-2000; 2000WO-US014941.
 PR 02-JUN-2000; 2000WO-US015264.
 PR 28-JUL-2000; 2000WO-US020710.
 PR 24-AUG-2000; 2000WO-US023278.
 PR 01-DEC-2000; 2000WO-US032678.
 PR 20-DEC-2000; 2000WO-US034956.
 PR 28-FEB-2001; 2001WO-US006520.
 PR 22-MAR-2001; 2001WO-US009552.
 PR 25-MAY-2001; 2001WO-US017092.
 PR 01-JUN-2001; 2001WO-US017800.
 PR 20-JUN-2001; 2001WO-US019692.
 PR 29-JUN-2001; 2001WO-US021066.
 PR 09-JUL-2001; 2001WO-US021735.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Baker KP, Bolstein D, Desnoyers L, Eaton DL;
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
 PI Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan MJ;
 PI Kijavrin JJ, Kuo SS, Napier MA, Pan J, Paoletti NF, Roy MA, Shelton DL;
 PI Stewart TA, Tumas D, Williams PM, Wood WI;
 XX
 DR WPI: 2003-328660/31.
 DR N-PSDB; ACA63545.
 XX
 PT New secreted and transmembrane nucleic acids and polypeptides, designated
 PT as PRO, useful for treating inflammation, organ failure, atherosclerosis,
 PT cardiac injury, infertility, birth defects, premature aging, AIDS, or
 PT cancer.
 PT
 XX
 PS Claim 12, Fig 39; 453pp; English.
 XX
 CC The invention describes an isolated nucleic acid (1) comprising, or which
 CC is at least 80 % sequence identity to, or the full-length coding sequence
 CC of, any of 118 300-2100 nucleotide sequences, which encodes its
 CC corresponding PRO polypeptide selected from 118 100-700 amino acid
 CC sequences, all given in the specification. The nucleic acids and

polypeptides are useful for treating inflammatory diseases, organ failure, atherosclerosis, cardiac injury, infertility, birth defects, premature aging, AIDS, cancer, or diabetic complications. The nucleic acids are useful as hybridisation probes, in chromosome and gene mapping, and in generating antisense RNA or DNA. The polypeptides are useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both are useful in tissue typing. This is the amino acid sequence of a novel human secreted and transmembrane PRO polypeptide

Sequence 730 AA:

Query Match 100.0%; Score 3843; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MGVCQRTAPWKEKSQLERALLGFRKGGSGMFASGMNQTVPIEAGSMALLPLILL 60
DB 1 MGVCQRTAPWKEKSQLERALLGFRKGGSGMFASGMNQTVPIEAGSMALLPLILL 60
QY 61 PLLIKLHLMPQLRWLPADLAFAYRALCCRRALRARAALAAAADPEGEGCSIAWRLAE 120
DB 61 PLLIKLHLMPQLRWLPADLAFAYRALCCRRALRARAALAAAADPEGEGCSIAWRLAE 120
QY 61 PLLIKLHLMPQLRWLPADLAFAYRALCCRRALRARAALAAAADPEGEGCSIAWRLAE 120
DB 61 PLLIKLHLMPQLRWLPADLAFAYRALCCRRALRARAALAAAADPEGEGCSIAWRLAE 120
QY 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
DB 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
QY 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
DB 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
QY 181 APGAGDAAGSAGAEFAGDGAARGGAAAPLSPGATVALLPAGPEFLIMFGIAKAGLR 240
DB 181 APGAGDAAGSAGAEFAGDGAARGGAAAPLSPGATVALLPAGPEFLIMFGIAKAGLR 240
QY 181 APGAGDAAGSAGAEFAGDGAARGGAAAPLSPGATVALLPAGPEFLIMFGIAKAGLR 240
DB 181 APGAGDAAGSAGAEFAGDGAARGGAAAPLSPGATVALLPAGPEFLIMFGIAKAGLR 240
QY 241 TAFVPTALRRGPLHCHLRSCGARALVLAPELSELEPDLPLRMAGHLMWAGSTPAG 300
DB 241 TAFVPTALRRGPLHCHLRSCGARALVLAPELSELEPDLPLRMAGHLMWAGSTPAG 300
QY 241 TAFVPTALRRGPLHCHLRSCGARALVLAPELSELEPDLPLRMAGHLMWAGSTPAG 300
DB 241 TAFVPTALRRGPLHCHLRSCGARALVLAPELSELEPDLPLRMAGHLMWAGSTPAG 300
QY 301 ISDLIAEVSAREVGPVPGYSSPISITDCLYIFTSQTTGPKXARISHLILCOQGRYQ 360
DB 301 ISDLIAEVSAREVGPVPGYSSPISITDCLYIFTSQTTGPKXARISHLILCOQGRYQ 360
QY 301 ISDLIAEVSAREVGPVPGYSSPISITDCLYIFTSQTTGPKXARISHLILCOQGRYQ 360
DB 301 ISDLIAEVSAREVGPVPGYSSPISITDCLYIFTSQTTGPKXARISHLILCOQGRYQ 360
QY 361 LCGHQEDEVLYIALPLVHMSGSLGIVCGMGIGATVVKSRFSKQFVEDCCQHRVYFQ 420
DB 361 LCGHQEDEVLYIALPLVHMSGSLGIVCGMGIGATVVKSRFSKQFVEDCCQHRVYFQ 420
QY 361 LCGHQEDEVLYIALPLVHMSGSLGIVCGMGIGATVVKSRFSKQFVEDCCQHRVYFQ 420
DB 361 LCGHQEDEVLYIALPLVHMSGSLGIVCGMGIGATVVKSRFSKQFVEDCCQHRVYFQ 420
QY 421 YIGELCRVLVNOPEKARHGKVRILAVSGSLRPTWSEFVRFGLQVLETYGHTGEGVA 480
DB 421 YIGELCRVLVNOPEKARHGKVRILAVSGSLRPTWSEFVRFGLQVLETYGHTGEGVA 480
QY 421 YIGELCRVLVNOPEKARHGKVRILAVSGSLRPTWSEFVRFGLQVLETYGHTGEGVA 480
DB 421 YIGELCRVLVNOPEKARHGKVRILAVSGSLRPTWSEFVRFGLQVLETYGHTGEGVA 480
QY 481 TINVTGQGAAGRASWLYKHIFPESLIHYDVTGTGPIIDPOGHCMATSPGEGILLAVPS 540
DB 481 TINVTGQGAAGRASWLYKHIFPESLIHYDVTGTGPIIDPOGHCMATSPGEGILLAVPS 540
QY 481 TINVTGQGAAGRASWLYKHIFPESLIHYDVTGTGPIIDPOGHCMATSPGEGILLAVPS 540
DB 481 TINVTGQGAAGRASWLYKHIFPESLIHYDVTGTGPIIDPOGHCMATSPGEGILLAVPS 540
QY 541 QOSPGLVAGGPELAAQGLKLDVFRPBDVFNFTGDLVCDQGRFRFDRIGDTFRMKGE 600
DB 541 QOSPGLVAGGPELAAQGLKLDVFRPBDVFNFTGDLVCDQGRFRFDRIGDTFRMKGE 600
QY 541 QOSPGLVAGGPELAAQGLKLDVFRPBDVFNFTGDLVCDQGRFRFDRIGDTFRMKGE 600
DB 541 QOSPGLVAGGPELAAQGLKLDVFRPBDVFNFTGDLVCDQGRFRFDRIGDTFRMKGE 600
QY 601 NVATTEVAEVEFALDFLOEVNVYGVTVPGHGRAGMAALVLRPPHLDLMQLYTHVSENL 660
DB 601 NVATTEVAEVEFALDFLOEVNVYGVTVPGHGRAGMAALVLRPPHLDLMQLYTHVSENL 660
QY 601 NVATTEVAEVEFALDFLOEVNVYGVTVPGHGRAGMAALVLRPPHLDLMQLYTHVSENL 660
DB 601 NVATTEVAEVEFALDFLOEVNVYGVTVPGHGRAGMAALVLRPPHLDLMQLYTHVSENL 660
QY 661 PPYAPRRLLOESLATTETFRQCKVRMANGFPDSTLSDLYLDDQVGVLYLTARY 720
DB 661 PPYAPRRLLOESLATTETFRQCKVRMANGFPDSTLSDLYLDDQVGVLYLTARY 720
QY 661 PPYAPRRLLOESLATTETFRQCKVRMANGFPDSTLSDLYLDDQVGVLYLTARY 720
DB 661 PPYAPRRLLOESLATTETFRQCKVRMANGFPDSTLSDLYLDDQVGVLYLTARY 720
QY 721 SALLAGNLR 730
DB 721 SALLAGNLR 730
QY 721 SALLAGNLR 730
DB 721 SALLAGNLR 730

```

RESULT 8
ADA84115
ID ADA84115 standard; protein; 730 AA.
XX
AC ADA84115;

```

XX 20-NOV-2003 (first entry)
XX Human POW124 protein.
DE human; marker; expressed sequence tag; EST; arabisopsis; tumour;
KW stress-induced phenotype; hypertonic stress; colon cancer; immunogen;
KW vaccine.
XX Homo sapiens.
XX NC2002103028-A2.
XX 27-DEC-2002.
XX 30-MAY-2002; 2002MO-1B004189.
XX 30-MAY-2001; 2001US-0293999P.
XX 22-OCT-2001; 2001US-0330457P.
XX 19-FEB-2002; 2002US-0357144P.
XX (BIOM-) BIONMEDICAL CENT.
XX Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;
XX WPI; 2003-175241/17.
XX N-PSDB; ADA84114.
XX Determining if a nucleic acid is a marker for a phenotype/cell type of
XX interest, by global comparison of expressed sequence tags known to be
XX expressed in the phenotype/cell type with all ESTs expressed in normal
XX tissue.
XX Claim 29; Page 491-494; 516pp; English.
XX The invention relates to a novel method for determining if a nucleic acid
XX is a marker for a predetermined phenotype/cell type of interest from a
XX biological species. The method comprises performing a global comparison
XX of a group of expressed sequence tags (ESTs) known to be expressed in the
XX phenotype/cell type of interest with all ESTs expressed in normal tissue
XX in order to identify ESTs that are preferentially expressed in the
XX phenotype/cell type of interest. A method of the invention is useful for
XX determining whether a nucleic acid is a marker for a predetermined
XX phenotype or cell type of interest from a biological species, preferably
XX Arabidopsis or human. The cell type of interest is an abnormal cell such
XX as a tumour cell, and the predetermined phenotype is a stress-induced
XX phenotype such as hypertonic stress or high salt conditions. A method
XX of the invention is also useful for determining the progression of colon
XX cancer in a human, for detecting a tumour cell, and for regulating or
XX preventing the growth of a tumour cell. An antibody of the invention is
XX useful for detecting the absence or presence of peptides encoded by
XX tumour-associated markers. A polypeptide of the invention is useful as an
XX immunogen for vaccinating an animal. The present sequence represents a
XX tumour-associated antigen of the invention.
XX
XX Sequence 730 AA:
QY 1 MGVCQRTAPWKEKSQLERALLGFRKGGSGMFASGMNQTVPIEAGSMALLPLILL 60
DB 1 MGVCQRTAPWKEKSQLERALLGFRKGGSGMFASGMNQTVPIEAGSMALLPLILL 60
QY 61 PLLIKLHLMPQLRWLPADLAFAYRALCCRRALRARAALAAAADPEGEGCSIAWRLAE 120
DB 61 PLLIKLHLMPQLRWLPADLAFAYRALCCRRALRARAALAAAADPEGEGCSIAWRLAE 120
QY 61 PLLIKLHLMPQLRWLPADLAFAYRALCCRRALRARAALAAAADPEGEGCSIAWRLAE 120
DB 61 PLLIKLHLMPQLRWLPADLAFAYRALCCRRALRARAALAAAADPEGEGCSIAWRLAE 120
QY 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
DB 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
QY 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180
DB 121 LAOQRAAHTFLIHGSRFRFSYSEARESNRARAFLRALGMDMGPDGDSGEGSAGEGERA 180

```


PR 11-MAR-1998; 98US-0077649P.
 PR 12-MAR-1998; 98US-0077791P.
 PR 13-MAR-1998; 98US-0078004P.
 PR 17-MAR-1998; 98US-00040220.
 PR 20-MAR-1998; 98US-0078886P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079254P.
 PR 26-MAR-1998; 98US-0079666P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079689P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 26-JUN-1998; 98US-00105413.
 PR 07-OCT-1998; 98US-0015697P.
 PR 07-OCT-1998; 98US-0021141.
 PR 02-NOV-1998; 98US-00184218.
 PR 06-NOV-1998; 98US-00187368.
 PR 20-NOV-1998; 98US-00202054.
 PR 07-DEC-1998; 98US-00202054.
 PR 22-DEC-1998; 98US-00218517.
 PR 05-JAN-1999; 98US-00000106.
 PR 05-MAR-1999; 98US-0025446P.
 PR 08-MAR-1999; 98US-00050502.
 PR 10-MAR-1999; 98US-00265686.
 PR 10-MAR-1999; 98US-00050519P.
 PR 12-APR-1999; 98US-00284291.
 PR 14-MAY-1999; 98US-00311833.
 PR 14-MAY-1999; 98US-00107733.
 PR 02-JUN-1999; 98US-00012252.
 PR 25-AUG-1999; 98US-00380137.
 PR 25-AUG-1999; 98US-00380138.
 PR 25-AUG-1999; 98US-00380142.
 PR 30-NOV-1999; 98US-00283313.
 PR 02-DEC-1999; 98US-00285551.
 PR 02-DEC-1999; 98US-00285551.
 PR 16-DEC-1999; 98US-00300095.
 PR 30-DEC-1999; 98US-0031243.
 PR 30-DEC-1999; 98US-0031274.
 PR 05-JAN-2000; 2000US-0000219.
 PR 06-JAN-2000; 2000US-0000227.
 PR 11-FEB-2000; 2000US-0000376.
 PR 11-FEB-2000; 2000US-00003565.
 PR 18-FEB-2000; 2000US-0004341.
 PR 24-FEB-2000; 2000US-0005004.
 PR 02-MAR-2000; 2000US-0005841.
 PR 10-MAR-2000; 2000US-0006319.
 PR 21-MAR-2000; 2000US-0007532.
 PR 30-MAR-2000; 2000US-0008439.
 PR 17-MAY-2000; 2000US-0013705.
 PR 22-MAY-2000; 2000US-0014042.
 PR 30-MAY-2000; 2000US-0014941.
 PR 02-JUN-2000; 2000US-0015264.
 PR 28-JUL-2000; 2000US-0020710.
 PR 24-AUG-2000; 2000US-0023328.
 PR 08-NOV-2000; 2000US-00709238.
 PR 27-NOV-2000; 2000US-00723749.
 PR 01-DEC-2000; 2000US-0082678.
 PR 20-DEC-2000; 2000US-00747255.
 PR 20-DEC-2000; 2000US-0034956.
 PR 28-FEB-2001; 2001US-0006520.
 PR 22-MAR-2001; 2001US-00816744.
 PR 22-MAR-2001; 2001US-00819920.
 PR 22-MAR-2001; 2001US-00099552.
 PR 10-MAY-2001; 2001US-00854208.
 PR 10-MAY-2001; 2001US-00854280.
 PR 25-MAY-2001; 2001US-00871092.
 PR 01-JUN-2001; 2001US-00872035.
 PR 01-JUN-2001; 2001US-0017800.

PR 05-JUN-2001; 2001US-00874503.
 PR 14-JUN-2001; 2001US-00882636.
 PR 19-JUN-2001; 2001US-00886342.
 PR 20-JUN-2001; 2001US-0019692.
 PR 23-JUN-2001; 2001US-0021066.
 PR 09-JUL-2001; 2001US-0021735.
 PR 30-JUL-2001; 2001US-00918585.
 XX
 XX (GENTH) GENENTECH INC.
 PA
 XX Ashkenazi A, Baker KP, Botstein D, Desnoyers J, Eaton D;
 PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
 PI Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ;
 PI Kijavrin IU, Kuo SS, Napier MA, Pan J, Paoletti NF, Roy MA, Shelton DL;
 PI Stewart TA, Tumas D, Williams PM, Wood WI;
 XX
 XX WPI; 2003-288163/28.
 DR N-PSDB; ABX92349.
 XX
 PT Novel secreted and transmembrane polypeptides and polynucleotides
 PT encoding them useful for treating cancer, kidney diseases, bone,
 PT cartilage disorders and immune deficiencies.
 PS
 PS Claim 12; Fig 39; 459pp; English.
 XX
 CC The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The PRO
 CC polypeptides are secreted and transmembrane proteins. The PRO
 CC polypeptides are useful for detecting other PRO polypeptides, for linking
 CC bioactive molecules to cells expressing PRO polypeptides, for modulating
 CC biological activities of cells expressing PRO polypeptides, and for for
 CC identifying agonists or antagonists. The bioactive molecule may be a
 CC toxin, radiolabel or antibody, and causes apoptosis or death of the cell.
 CC The PRO polypeptides are useful for treating immune disorders, diabetes
 CC or hyper- or hypo-insulinaemia, cardiac insufficiency, nervous system
 CC disorders, kidney disorders, bone and cartilage disorders or arthritis,
 CC tumours, and wound healing. The polynucleotide sequences encoding PRO
 CC polypeptides are useful as hybridisation probes, in chromosome and gene
 CC mapping, in the generation of antisense RNA and DNA, in the preparation
 CC of PRO polypeptides, for generating transgenic animals or knockout
 CC animals, for the genetic analysis of individuals with genetic disorders,
 CC and in gene therapy. ABUS1071-ABUS16164 represent the human PRO
 CC polypeptides of the invention. Note: The sequence data for this patent
 CC was obtained in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/psipdsidentry.html
 XX
 SO Sequence 730 AA;
 Query Match 100.0%; Score 3843; DB 6; Length 730;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGVCORTRAPWKEKSLERAAALGFRKSGSGMFASGWNQTVPIEAGSMALILLPLILL 60
 DB 1 MGVCORTRAPWKEKSLERAAALGFRKSGSGMFASGWNQTVPIEAGSMALILLPLILL 60
 QY 61 PLILFLHIMPOLRMLPADLAFAVRALCCRRALRARAALAAADDEPGGCSLAWRLAE 120
 DB 61 PLILFLHIMPOLRMLPADLAFAVRALCCRRALRARAALAAADDEPGGCSLAWRLAE 120
 QY 121 LAQGRRAHTFLIHGSRFYSSEARESNRAAPAFALGMDMPGPGGSGEGSAGEGERA 180
 DB 121 LAQGRRAHTFLIHGSRFYSSEARESNRAAPAFALGMDMPGPGGSGEGSAGEGERA 180
 QY 121 LAQGRRAHTFLIHGSRFYSSEARESNRAAPAFALGMDMPGPGGSGEGSAGEGERA 180
 DB 121 LAQGRRAHTFLIHGSRFYSSEARESNRAAPAFALGMDMPGPGGSGEGSAGEGERA 180
 QY 181 APGAGDAAAGSGAFAGDGAARGGGAAAPLSPGATVALLDPAGEPFLWFGAKAGR 240
 DB 181 APGAGDAAAGSGAFAGDGAARGGGAAAPLSPGATVALLDPAGEPFLWFGAKAGR 240
 QY 241 TAFVPTALRGRLHGLRSCGARALVLAPEPESLEPDI PARANGIHWAGPGTHPAG 300
 DB 241 TAFVPTALRGRLHGLRSCGARALVLAPEPESLEPDI PARANGIHWAGPGTHPAG 300
 QY 301 ISDLIAEVAEYDGPVGYLSSPOSITDTCLYIFTSGITGLPKAARISHLKIQCQGFYQ 360

30-JUL-2001; 2001US-00918585.

(GENTH) GENENTECH INC.

Ashkenazi AJ, Baker KP, Botstein D, Desnuciers L, Baton DL, Ferrara N, Flvairoff E, Fong S, Gao W, Gerder H, Gertlissen ME, Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KJ, Kijavlin IJ, Kuo SS, Napier MA, Par J, Paoni NF, Roy MA, Shelton DL, Stewart TA, Tumes D, Williams PM, Wood WJ; WPI, 2003-341189/32. N-PSDB; ACA66090.

New genes and secreted and transmembrane polypeptides (e.g. PRO337 or PRO1559), useful for treating or diagnosing e.g. cancers, atherosclerosis, infertility, stroke, encephalitis, hepatitis or multiple sclerosis in mammals.

Claim 12; Fig 39; 460pp; English.

The invention relates to a new isolated nucleic acid molecule comprising a sequence with at least 80% identity to: (a) a nucleotide encoding any of 94 PRO polypeptides whose sequences are fully defined in the specification; or (b) any of 94 nucleotide sequences fully defined in the specification; or the full length coding sequence of any these 94 nucleotide sequences. Also included are an isolated PRO polypeptide scoring at least 80% positives when compared to any of the PRO polypeptide sequences cited above (or an isolated PRO polypeptide having at least 80% amino acid sequence identity to: (a) an amino acid sequence encoded by the nucleotide deposited with ATCC numbers listed in the specification; (b) the PRO polypeptide, lacking its associated signal peptide; or (c) an extracellular domain of the PRO polypeptide, with or lacking its associated signal peptide), a vector comprising the nucleic acid molecule, a host cell comprising the vector (and producing a PRO polypeptide), a chimeric molecule comprising the PRO polypeptide fused to a heterologous amino acid sequence and an anti-PRO antibody. The PRO polypeptides or polynucleotides are useful as pharmaceuticals, diagnostics, biosensors or bioreactors. These are particularly useful for detecting or creating e.g. malignancies or cancers (e.g. ovarian cancer, colorectal cancer, sarcoma, leukaemia or lymphoma), inflammatory disease, necrosis, atherosclerosis, infertility, premature aging, psoriasis, inflammation, hepatitis, or multiple sclerosis in mammals. The PRO polypeptides are useful in drug screening, particularly as targets for therapeutic intervention in these diseases, and in the diagnostic determination of the presence of these diseases. The PRO polypeptides are also useful as molecular weight markers, or for chromosome identification. The PRO genes are useful as hybridisation probes, or for screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene therapy, particularly for replacing a defective gene. The present sequence represents a PRO polypeptide

Query Match 100.0%; Score 3843; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MGVCORTRAPWKEKQLEPAALGFRKGGSGMFGSMNQTVPHEEGSMALLLPILL 60
1 MGVCORTRAPWKEKQLEPAALGFRKGGSGMFGSMNQTVPHEEGSMALLLPILL 60
61 PILLIKLHMPOLRLPADLAFAYALCKKALRRARALAAAADPEEGGCSLAMRLAE 120
61 PILLIKLHMPOLRLPADLAFAYALCKKALRRARALAAAADPEEGGCSLAMRLAE 120
121 LAOQRAAHTFLHSGRRRSYSFAERESNRARAFRALGMDWGPDDGSGSGSAGEGERA 180
121 LAOQRAAHTFLHSGRRRSYSFAERESNRARAFRALGMDWGPDDGSGSGSAGEGERA 180
181 AFGADDAAGSAEFAEGDGAARCGAAPTSPGATVALLPAGPEFLMIFGAKAGLR 240
181 AFGADDAAGSAEFAEGDGAARCGAAPTSPGATVALLPAGPEFLMIFGAKAGLR 240

241 TAFVPTALRGPLHLCLRSCGARALYLAPEPLESEPPDLPALRAMGLHMAAGPGRHAPAG 300
241 TAFVPTALRGPLHLCLRSCGARALYLAPEPLESEPPDLPALRAMGLHMAAGPGRHAPAG 300
301 ISDLAEVSAEVDGPPVPGILSSPQSTITDCLVITFTSGITGLPKPARISHUKLQCGEYQ 360
301 ISDLAEVSAEVDGPPVPGILSSPQSTITDCLVITFTSGITGLPKPARISHUKLQCGEYQ 360
361 LCGVHEDVITYALPETHMSSGLIGVGMGIGATVLSKXSAGOFMEDCOOHRVTVQ 420
361 LCGVHEDVITYALPETHMSSGLIGVGMGIGATVLSKXSAGOFMEDCOOHRVTVQ 420
421 YIGELCRYLVNPPSPKAERGHKRLAVSGSLPDTWERFRRRFGPLQVLETYGLTEGNA 480
421 YIGELCRYLVNPPSPKAERGHKRLAVSGSLPDTWERFRRRFGPLQVLETYGLTEGNA 480
481 TINYTGQRAGVGAASLKYHIFPFSLIRVDVTGTEIRDPQCHMATSPGEGLLVAPVS 540
481 TINYTGQRAGVGAASLKYHIFPFSLIRVDVTGTEIRDPQCHMATSPGEGLLVAPVS 540
541 QQSPFLGVAGPELAAQGLKLVFREGDVFNTGDLVCDQGFRLFHRTDTRMKGE 600
541 QQSPFLGVAGPELAAQGLKLVFREGDVFNTGDLVCDQGFRLFHRTDTRMKGE 600
601 NVATTEVAEVEFALDPLQEVNNTYGVTVPHGEGRAGMAALVLRPPHALDMQLYTHSENL 660
601 NVATTEVAEVEFALDPLQEVNNTYGVTVPHGEGRAGMAALVLRPPHALDMQLYTHSENL 660
661 PPVAPRPFRLQBSLATTEFPKQKVRMANEGFDSTLSDPPLYVLDQAVGYLPTTARY 720
661 PPVAPRPFRLQBSLATTEFPKQKVRMANEGFDSTLSDPPLYVLDQAVGYLPTTARY 720
721 SALLAGNLRI 730
721 SALLAGNLRI 730

RESULT 12
ADA24641
ID ADA24641 standard; protein; 730 AA.
XX
ADA24641;
XX
20-NOV-2003 (first entry)
XX
XX
DE Novel human secreted and transmembrane protein PRO703.
XX
XX Human; secreted and transmembrane protein; PRO; tissue typing;
XX chromosome identification; vaccine; cancer; retinal disorder;
XX sports-related joint disorder; osteoarthritis; rheumatoid arthritis;
XX wound healing; obesity; diabetes; hearing loss;
XX cardiac insufficiency disorder; kidney disorder; nervous system disorder;
XX haemoglobin associated disorder.
XX
OS Homo sapiens.
PN US2003050241-A1.
PD 13-MAR-2003.
XX
XX
PF 16-OCT-2001; 2001US-00978564.
XX
XX 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.

PR 20-MAR-1998; 98US-0078866P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 25-MAR-1998; 98US-0079294P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079688P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 27-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 31-MAR-1998; 98US-0080105P.
 PR 31-MAR-1998; 98US-0080107P.
 PR 31-MAR-1998; 98US-0080165P.
 PR 31-MAR-1998; 98US-0080194P.
 PR 01-APR-1998; 98US-0080327P.
 PR 01-APR-1998; 98US-0080328P.
 PR 01-APR-1998; 98US-0080333P.
 PR 01-APR-1998; 98US-0080334P.
 PR 08-APR-1998; 98US-0081049P.
 PR 08-APR-1998; 98US-0081070P.
 PR 08-APR-1998; 98US-0081071P.
 PR 09-APR-1998; 98US-0081195P.
 PR 09-APR-1998; 98US-0081203P.
 PR 09-APR-1998; 98US-0081223P.
 PR 15-APR-1998; 98US-0081817P.
 PR 15-APR-1998; 98US-0081818P.
 PR 15-APR-1998; 98US-0081838P.
 PR 15-APR-1998; 98US-0081952P.
 PR 15-APR-1998; 98US-0081955P.
 PR 21-APR-1998; 98US-0082568P.
 PR 21-APR-1998; 98US-0082569P.
 PR 22-APR-1998; 98US-0082700P.
 PR 22-APR-1998; 98US-0082704P.
 PR 22-APR-1998; 98US-0082797P.
 PR 22-APR-1998; 98US-0082804P.
 PR 23-APR-1998; 98US-0082796P.
 PR 27-APR-1998; 98US-0083336P.
 PR 28-APR-1998; 98US-0083322P.
 PR 29-APR-1998; 98US-0083392P.
 PR 29-APR-1998; 98US-0083455P.
 PR 29-APR-1998; 98US-0083496P.
 PR 29-APR-1998; 98US-0083499P.
 PR 29-APR-1998; 98US-0083500P.
 PR 29-APR-1998; 98US-0083545P.
 PR 29-APR-1998; 98US-0083554P.
 PR 29-APR-1998; 98US-0083558P.
 PR 29-APR-1998; 98US-0083559P.
 PR 30-APR-1998; 98US-0083742P.
 PR 05-MAY-1998; 98US-0084366P.
 PR 06-MAY-1998; 98US-0084414P.
 PR 06-MAY-1998; 98US-0084441P.
 PR 07-MAY-1998; 98US-0084598P.
 PR 07-MAY-1998; 98US-0084600P.
 PR 07-MAY-1998; 98US-0084627P.
 PR 07-MAY-1998; 98US-0084637P.
 PR 07-MAY-1998; 98US-0084639P.
 PR 07-MAY-1998; 98US-0084640P.
 PR 07-MAY-1998; 98US-0084643P.
 PR 13-MAY-1998; 98US-0085323P.
 PR 13-MAY-1998; 98US-0085338P.
 PR 13-MAY-1998; 98US-0085339P.
 PR 15-MAY-1998; 98US-0085573P.
 PR 15-MAY-1998; 98US-0085579P.
 PR 15-MAY-1998; 98US-0085580P.
 PR 15-MAY-1998; 98US-0085582P.
 PR 15-MAY-1998; 98US-0085689P.
 PR 15-MAY-1998; 98US-0085697P.
 PR 15-MAY-1998; 98US-0085700P.
 PR 15-MAY-1998; 98US-0085704P.
 PR 18-MAY-1998; 98US-0086023P.

PR 22-MAY-1998; 98US-0086392P.
 PR 22-MAY-1998; 98US-0086414P.
 PR 22-MAY-1998; 98US-0086430P.
 PR 22-MAY-1998; 98US-0086486P.
 PR 28-MAY-1998; 98US-0087098P.
 PR 28-MAY-1998; 98US-0087106P.
 PR 28-MAY-1998; 98US-0087208P.
 PR 26-JUN-1998; 98US-0090863P.
 PR 26-JUN-1998; 98US-0091010P.
 PR 01-JUL-1998; 98US-0091359P.
 PR 30-JUL-1998; 98US-0094651P.
 PR 11-SEP-1998; 98US-0100038P.
 PR 07-OCT-1998; 98US-0109214P.
 PR 20-NOV-1998; 98US-0109304P.
 PR 20-NOV-1998; 98US-0123957P.
 PR 22-DEC-1998; 98US-0113296P.
 PR 22-DEC-1998; 98US-0113621P.
 PR 05-JAN-1999; 99MO-US000106.
 PR 08-MAR-1999; 99MO-US005028.
 PR 10-MAR-1999; 99MO-US005190.
 PR 12-MAR-1999; 99US-0123957P.
 PR 29-MAR-1999; 99US-0126773P.
 PR 21-APR-1999; 99US-0130232P.
 PR 26-APR-1999; 99US-0131022P.
 PR 28-APR-1999; 99US-0131445P.
 PR 14-MAY-1999; 99US-0134287P.
 PR 14-MAY-1999; 99MO-US010733.
 PR 02-JUN-1999; 99MO-US01252.
 PR 16-JUN-1999; 99US-0139557P.
 PR 23-JUN-1999; 99US-0141037P.
 PR 07-JUL-1999; 99US-0142680P.
 PR 26-JUL-1999; 99US-0145598P.
 PR 28-JUL-1999; 99US-0146222P.
 PR 29-OCT-1999; 99US-0162506P.
 PR 30-NOV-1999; 99MO-US028313.
 PR 02-DEC-1999; 99MO-US028551.
 PR 02-DEC-1999; 99MO-US028565.
 PR 16-DEC-1999; 99MO-US030095.
 PR 30-DEC-1999; 99MO-US031274.
 PR 30-DEC-1999; 99MO-US031283.
 PR 05-JAN-2000; 2000MO-US000219.
 PR 06-JAN-2000; 2000MO-US000277.
 PR 06-JAN-2000; 2000MO-US000376.
 PR 11-FEB-2000; 2000MO-US000365.
 PR 18-FEB-2000; 2000MO-US004341.
 PR 24-FEB-2000; 2000MO-US005004.
 PR 02-MAR-2000; 2000MO-US005841.
 PR 10-MAR-2000; 2000MO-US006319.
 PR 21-MAR-2000; 2000MO-US007532.
 PR 30-MAR-2000; 2000MO-US008439.
 PR 17-MAY-2000; 2000MO-US013705.
 PR 22-MAY-2000; 2000MO-US014042.
 PR 30-MAY-2000; 2000MO-US014941.
 PR 02-JUN-2000; 2000MO-US015264.
 PR 28-JUL-2000; 2000MO-US020710.
 PR 24-AUG-2000; 2000MO-US023328.
 PR 01-DEC-2000; 2000MO-US032678.
 PR 20-DEC-2000; 2000MO-US034956.
 PR 28-FEB-2001; 2001MO-US006520.
 PR 22-MAR-2001; 2001MO-US009552.
 PR 25-MAY-2001; 2001MO-US017092.
 PR 01-JUN-2001; 2001MO-US017800.
 PR 20-JUN-2001; 2001MO-US019692.
 PR 29-JUN-2001; 2001MO-US021066.
 PR 09-JUL-2001; 2001MO-US021735.
 PR 30-JUL-2001; 2001US-00918585.
 (GETH) GENENTECH INC.
 XX
 PA Ashkenazi AJ, Baker KP, Bolstein D, Desnoyers L, Eaton DL;
 PI Perrata N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
 PI Goddard A, Godowski FO, Grimaldi JC, Gurney AL, Hillian KU;
 PI Kijavlin IU, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;

PI Stewart TA, Tumas D, Williams PM, Wood WI;
 XX WPI; 2003-521814/49.
 DR N-PSDB; ADA24640.
 XX
 PT New isolated PRO polypeptides for example extracellular, secreted and
 PT membrane bound proteins, useful for modulating the biological activities
 PT of cells and for treating, for example diabetes, cancer, rheumatoid
 PT arthritis, and hearing loss.
 XX
 XX Claim 12; Fig 39; 461p; English.

CC The invention describes an isolated secreted and transmembrane (PRO)
 CC polypeptide (1). PRO337 polypeptide is useful for detecting PRO4993
 CC polypeptide in a sample, and vice versa. PRO725, PRO700 and PRO739 are
 CC useful for detecting PRO1559 polypeptide in a sample, and PRO1559 is
 CC useful for detecting PRO725, PRO700 and PRO739 in a sample. PRO4993 is
 CC useful for linking a bioactive molecule to a cell expressing a PRO337
 CC polypeptide, and PRO337 is useful for linking a bioactive molecule to a
 CC cell expressing a PRO4993 polypeptide. PRO1559 is useful for linking a
 CC bioactive molecule to a cell expressing a PRO735, PRO700 and PRO739

Query Match

100.0%; Score 3843; DB 6; Length 730;

Best Local Similarity 100.0%; Pred. No. 0;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTRAPMKEKSLERPAALGPRKSGGSMFASGNNOTVPIEEAGSYAALLPILLL 60
 DB 1 MGVCORTRAPMKEKSLERPAALGPRKSGGSMFASGNNOTVPIEEAGSYAALLPILLL 60
 QY 61 PLLLLKLHLMPQLRMLPADLAFAYRALCCKRALRRALAAAADDEGGGCSLAWRLAE 120
 DB 61 PLLLLKLHLMPQLRMLPADLAFAYRALCCKRALRRALAAAADDEGGGCSLAWRLAE 120
 QY 121 LMOQRAHTFLHSGRRSEYSAERBSNPARAFPLAGMDGPGGWSGSGSAGEGERRA 180
 DB 121 LMOQRAHTFLHSGRRSEYSAERBSNPARAFPLAGMDGPGGWSGSGSAGEGERRA 180
 QY 181 AFGAGDAAAGSAGAEFAGDGAARCGAAPLSPGATVALLPAGPEFLMFWGLAKAGLR 240
 DB 181 AFGAGDAAAGSAGAEFAGDGAARCGAAPLSPGATVALLPAGPEFLMFWGLAKAGLR 240
 QY 241 TAFVPAALRGPLHLCRCGARALVLAPELIESLPPLPALRANGHLMAAGGTHAG 300
 DB 241 TAFVPAALRGPLHLCRCGARALVLAPELIESLPPLPALRANGHLMAAGGTHAG 300
 QY 301 ISDLAEVSAEVDGPVPGYLSPOSITDCLYIFTSGETGLPKAARISHLKLQCGFYQ 360
 DB 301 ISDLAEVSAEVDGPVPGYLSPOSITDCLYIFTSGETGLPKAARISHLKLQCGFYQ 360
 QY 361 LCGVHOEDVLYIALPLYNHSGSLGIVGCMGIGATVILKSFSAQFWEDECQHRVTFQ 420
 DB 361 LCGVHOEDVLYIALPLYNHSGSLGIVGCMGIGATVILKSFSAQFWEDECQHRVTFQ 420
 QY 421 YIGELCRVYNOPPSAERGHKVRILAVSGLRPDWFERFVRFGLOVLETYGLTEGNA 480
 DB 421 YIGELCRVYNOPPSAERGHKVRILAVSGLRPDWFERFVRFGLOVLETYGLTEGNA 480
 QY 481 TINYTGORGAVRASWLYKHIFPSLIRYDVTGGPIRDPOCHMATSPPGEGLLVAVS 540
 DB 481 TINYTGORGAVRASWLYKHIFPSLIRYDVTGGPIRDPOCHMATSPPGEGLLVAVS 540
 QY 541 QOSPFLGAGGPELNOGKLLKDVFRPGVFFNTGILLVCDQOGFLRFHDFRGTFRMKE 600
 DB 541 QOSPFLGAGGPELNOGKLLKDVFRPGVFFNTGILLVCDQOGFLRFHDFRGTFRMKE 600
 QY 601 NVATTEVAEVPALDFLOEVNYYGTVPGHEGRAGAAALVLRPPALDLMOLYTHVENL 660
 DB 601 NVATTEVAEVPALDFLOEVNYYGTVPGHEGRAGAAALVLRPPALDLMOLYTHVENL 660
 QY 661 PPIARPRFLRLOESLATTSTTFKQOKVRANSGFDSITSDPLVYLDQVAGALPLTARY 720
 DB 661 PPIARPRFLRLOESLATTSTTFKQOKVRANSGFDSITSDPLVYLDQVAGALPLTARY 720

QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730
 RESULT 13
 ABO19656
 ID ABO19656 standard; protein; 730 AA.
 XX
 AC ABO19656;
 XX
 DT 08-SEP-2003 (first entry)
 XX
 DE Novel human secreted and transmembrane protein PRO703.
 XX
 KW Human, secreted and transmembrane protein; PRO; cell death; neuropathy;
 KW peripheral neuropathy; diabetic peripheral neuropathy;
 KW AIDS-associated neuropathy; Charcot-Marie-Tooth disease;
 KW Refsum's disease; Abetalipoproteinemia; Tangier disease;
 KW Krabbe's disease; Metachromatic leukodystrophy; Fabry's disease;
 KW Djerfene-Sottas syndrome; chromosome mapping; gene mapping; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN US2003050240-A1.
 XX
 PD 13-MAR-2003.
 XX
 PF 16-OCT-2001; 2001US-00978403.
 XX
 PR 17-OCT-1997; 97US-0062250P.
 PR 03-NOV-1997; 97US-0064249P.
 PR 13-NOV-1997; 97US-0065311P.
 PR 21-NOV-1997; 97US-0066364P.
 PR 10-MAR-1998; 98US-0077450P.
 PR 11-MAR-1998; 98US-0077532P.
 PR 11-MAR-1998; 98US-0077641P.
 PR 11-MAR-1998; 98US-0077649P.
 PR 12-MAR-1998; 98US-0077791P.
 PR 13-MAR-1998; 98US-0078004P.
 PR 20-MAR-1998; 98US-0078886P.
 PR 20-MAR-1998; 98US-0078910P.
 PR 20-MAR-1998; 98US-0078936P.
 PR 20-MAR-1998; 98US-0078939P.
 PR 26-MAR-1998; 98US-0079294P.
 PR 26-MAR-1998; 98US-0079656P.
 PR 27-MAR-1998; 98US-0079663P.
 PR 27-MAR-1998; 98US-0079664P.
 PR 27-MAR-1998; 98US-0079689P.
 PR 27-MAR-1998; 98US-0079728P.
 PR 27-MAR-1998; 98US-0079786P.
 PR 30-MAR-1998; 98US-0079920P.
 PR 30-MAR-1998; 98US-0079923P.
 PR 31-MAR-1998; 98US-0080105P.
 PR 31-MAR-1998; 98US-0080107P.
 PR 31-MAR-1998; 98US-0080165P.
 PR 31-MAR-1998; 98US-0080194P.
 PR 01-APR-1998; 98US-0080327P.
 PR 01-APR-1998; 98US-0080328P.
 PR 01-APR-1998; 98US-0080330P.
 PR 01-APR-1998; 98US-0080334P.
 PR 08-APR-1998; 98US-0081049P.
 PR 08-APR-1998; 98US-0081070P.
 PR 08-APR-1998; 98US-0081071P.
 PR 09-APR-1998; 98US-0081195P.
 PR 09-APR-1998; 98US-0081203P.
 PR 15-APR-1998; 98US-0081229P.
 PR 15-APR-1998; 98US-0081817P.
 PR 15-APR-1998; 98US-0081819P.
 PR 15-APR-1998; 98US-0081838P.
 PR 15-APR-1998; 98US-0081852P.
 PR 15-APR-1998; 98US-0081955P.

PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083335P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083546P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 05-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084441P.
PR 07-MAY-1998; 98US-0084598P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 15-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085573P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085689P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086032P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-0090663P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-01021141.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98US-0109304P.
PR 22-DEC-1998; 98US-0112262P.
PR 23-DEC-1998; 98US-0116212P.
PR 05-JAN-1999; 98US-0116212P.
PR 08-MAR-1999; 98US-0116212P.
PR 10-MAR-1999; 98US-0116212P.
PR 12-MAR-1999; 98US-0123957P.
PR 29-MAR-1999; 98US-0125773P.
PR 21-APR-1999; 98US-0130232P.
PR 26-APR-1999; 98US-0130422P.
PR 28-APR-1999; 98US-0130422P.
PR 14-MAY-1999; 98US-0134287P.
PR 14-MAY-1999; 98US-0134287P.
PR 02-JUN-1999; 98US-0134287P.
PR 16-JUN-1999; 98US-0134287P.
PR 23-JUN-1999; 98US-0141037P.
PR 07-JUL-1999; 98US-0142680P.
PR 26-JUL-1999; 98US-0145698P.

PR 28-JUL-1999; 98US-0146222P.
PR 28-OCT-1999; 98US-0162506P.
PR 30-NOV-1999; 98US-0162831P.
PR 02-DEC-1999; 98US-0162855P.
PR 02-DEC-1999; 98US-0162855P.
PR 16-DEC-1999; 98US-0162855P.
PR 30-DEC-1999; 98US-0162855P.
PR 30-DEC-1999; 98US-0162855P.
PR 05-JAN-2000; 98US-0162855P.
PR 06-JAN-2000; 98US-0162855P.
PR 11-FEB-2000; 98US-0162855P.
PR 11-FEB-2000; 98US-0162855P.
PR 18-FEB-2000; 98US-0162855P.
PR 24-FEB-2000; 98US-0162855P.
PR 02-MAR-2000; 98US-0162855P.
PR 10-MAR-2000; 98US-0162855P.
PR 21-MAR-2000; 98US-0162855P.
PR 30-MAR-2000; 98US-0162855P.
PR 17-MAY-2000; 98US-0162855P.
PR 22-MAY-2000; 98US-0162855P.
PR 30-MAY-2000; 98US-0162855P.
PR 02-JUN-2000; 98US-0162855P.
PR 28-JUL-2000; 98US-0162855P.
PR 28-AUG-2000; 98US-0162855P.
PR 01-DEC-2000; 98US-0162855P.
PR 20-DEC-2000; 98US-0162855P.
PR 28-FEB-2001; 98US-0162855P.
PR 22-MAR-2001; 98US-0162855P.
PR 25-MAY-2001; 98US-0162855P.
PR 01-JUN-2001; 98US-0162855P.
PR 20-JUN-2001; 98US-0162855P.
PR 29-JUN-2001; 98US-0162855P.
PR 09-JUL-2001; 98US-0162855P.
PR 30-JUL-2001; 98US-0162855P.
XX (GENTECH) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Deenoyers J, Eaton DJ,
XX Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME,
XX Goddard A, Godowski PJ, Grimaldi JC, Gurney AL, Hillan KU,
XX Kijavyn IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DJ,
XX Stewart TA, Tamas D, Williams PM, Wood WL,
XX WPI, 2003-503575/47.
XX N-PSDB; ACD29691.
XX Novel secreted and transmembrane polypeptide for modulating biological
XX activity of cell expressing the polypeptide, identifying agonists or
XX antagonists of polypeptide, and as molecular weight markers.
XX Claim 12; Fig 39; 459pp; English.
XX
XX The invention describes an isolated, secreted and transmembrane
XX polypeptide, termed PRO polypeptide (I). (I) is useful for detecting
XX PRO493, PRO337, PRO1559, PRO725, PRO700 or PRO739 polypeptide, and for
XX linking a bioactive molecule to a cell expressing the above polypeptide.
XX The bioactive molecule is a toxin, radiolabel or an antibody and causes
XX cell death. (I) is useful as therapeutic agent, in medical and industrial
XX applications e.g. for treating neuropathy, AIDS-associated neuropathy,
XX neuropathy, diabetic peripheral neuropathy, AIDS-associated neuropathy,
XX Charcot-Marie-Tooth disease, Refsum's disease, Abetalipoproteinemia,
XX Tangle disease, Krabbe's disease, Metachromatic leukodystrophy, Fabry's
XX
Query Match 100.0%; Score 3843; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVCORTPAFPMKKSQLEERAAALCFRKGSGMFAAGNQTVPLEEGSMALLPPLILL 60
DB 1 MGVCORTPAFPMKKSQLEERAAALCFRKGSGMFAAGNQTVPLEEGSMALLPPLILL 60
QY PLLILKLIHMPQILKMTPLADLAFVRLCCRAIRARALAAAADPPGEGCSLAWRLAE 120

Db 61 PLLIKLHIMPQLRLPADLAFVAVLALCKRALRRLALAAAAADPEEGCCSLAKRLAE 120
QY 121 LAQRAAHTLLIHGSRRTSYSAERESNRAPAFRLALGMDPGDGGSGSGAGEGERRA 180
Db 121 LAQRAAHTLLIHGSRRTSYSAERESNRAPAFRLALGMDPGDGGSGSGAGEGERRA 180
QY 181 ARGADGMAAGSGAEFAGGAGARGGAAAPLSPGATVLLLLPAGEPEFLMWFGAKXGLR 240
Db 181 ARGADGMAAGSGAEFAGGAGARGGAAAPLSPGATVLLLLPAGEPEFLMWFGAKXGLR 240
QY 241 TAFVPTALRGFLLHCLRSQCARALVLAPEFLSELEPDLPALRAMGLHLMAGPOTHPAG 300
Db 241 TAFVPTALRGFLLHCLRSQCARALVLAPEFLSELEPDLPALRAMGLHLMAGPOTHPAG 300
QY 301 ISDLIAEVAEVDGVPGLTSPGSIPTDCLYIFTSGTGPKARISHLKLLOCGGYQ 360
Db 301 ISDLIAEVAEVDGVPGLTSPGSIPTDCLYIFTSGTGPKARISHLKLLOCGGYQ 360
QY 361 LCGVHQBEDVLYALPLHNSGSLGLGVGCMGIGATVILKSKFSAGQFWECCQHRVTYQ 420
Db 361 LCGVHQBEDVLYALPLHNSGSLGLGVGCMGIGATVILKSKFSAGQFWECCQHRVTYQ 420
QY 421 YIGELCRVLYVNOPPSKAERGHKVRILAVSGSLRPDTWERFVRFGPLQVLETYGLTEGVVA 480
Db 421 YIGELCRVLYVNOPPSKAERGHKVRILAVSGSLRPDTWERFVRFGPLQVLETYGLTEGVVA 480
QY 481 TINYGORGAVRASMLYKHIFPFLIRYDTTGEPIRDPOGCMATSPGEGILLVAFVS 540
Db 481 TINYGORGAVRASMLYKHIFPFLIRYDTTGEPIRDPOGCMATSPGEGILLVAFVS 540
QY 541 QOSPFLGAVGPELLAQGKLKDVFRPGDVFNNTGDLVDDOGFLRFHDFRGTDFRWKGE 600
Db 541 QOSPFLGAVGPELLAQGKLKDVFRPGDVFNNTGDLVDDOGFLRFHDFRGTDFRWKGE 600
QY 601 NVATTEVAEVEFALDFLOEVNYYGVTVPGHEGRAGMALVLRPPALDLMQLYTHVSEN 660
Db 601 NVATTEVAEVEFALDFLOEVNYYGVTVPGHEGRAGMALVLRPPALDLMQLYTHVSEN 660
QY 661 PEYARPRFLRLOESLATTETFEKQKVRMANEGFDDSTLSDPLVYLDQAVAYLPPTTARY 720
Db 661 PEYARPRFLRLOESLATTETFEKQKVRMANEGFDDSTLSDPLVYLDQAVAYLPPTTARY 720
QY 721 SALLAGNLRI 730
Db 721 SALLAGNLRI 730

RESULT 14
ADA12302
ID ADA12302 standard, protein, 730 AA.
XX
AC ADA12302;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human secreted/transmembrane polypeptide PRO703.
XX
KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;
KW infertility; birth defect; premature aging; AIDS; cancer;
KW diabetic complication; tissue typing; human.
XX
OS Homo sapiens.
XX
PN US003055216-A1.
XX
PD 20-MAR-2003.
XX
PF 17-OCT-2001; 2001US-00978824.
XX
PR 21-MAY-1996; 96US-0018049P.
PR 17-OCT-1997; 97US-0062250B.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.

PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 12-MAR-1998; 98US-0077791P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-00804220.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078810P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079656P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080333P.
PR 01-APR-1998; 98US-0080334P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 09-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081203P.
PR 15-APR-1998; 98US-0081229P.
PR 15-APR-1998; 98US-0081817P.
PR 15-APR-1998; 98US-0081819P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0081852P.
PR 15-APR-1998; 98US-0081855P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083336P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083554P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084411P.
PR 07-MAY-1998; 98US-0084598P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 15-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085733P.
PR 15-MAY-1998; 98US-0085739P.

PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085689P.
PR 15-MAY-1998; 98US-0085697P.
PR 15-MAY-1998; 98US-0085700P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0093010P.
PR 01-JUL-1998; 98US-0093359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-01668978.
PR 07-OCT-1998; 98WO-US021141.
PR 02-NOV-1998; 98US-00184216.
PR 06-NOV-1998; 98US-00187358.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98WO-US024855.
PR 07-DEC-1998; 98US-00202054.
PR 22-DEC-1998; 98US-00218517.
PR 23-DEC-1998; 98US-0113296P.
PR 05-JAN-1999; 98WO-US000106.
PR 05-MAR-1999; 98US-00254455.
PR 08-MAR-1999; 98WO-US005028.
PR 10-MAR-1999; 98US-00265686.
PR 12-MAR-1999; 98US-00267213.
PR 12-MAR-1999; 98US-0123957P.
PR 29-MAR-1999; 98US-0125773P.
PR 12-APR-1999; 98US-00284231.
PR 21-APR-1999; 98US-0130232P.
PR 26-APR-1999; 98US-0131022P.
PR 28-APR-1999; 98US-0131445P.
PR 14-MAY-1999; 98US-00131832.
PR 14-MAY-1999; 98US-0134287P.
PR 02-JUN-1999; 98WO-US010733.
PR 16-JUN-1999; 98US-0139557P.
PR 23-JUN-1999; 98US-0141037P.
PR 07-JUL-1999; 98US-0142680P.
PR 26-JUL-1999; 98US-0145698P.
PR 28-JUL-1999; 98US-0146223P.
PR 25-AUG-1999; 98US-00380137.
PR 25-AUG-1999; 98US-00380138.
PR 25-AUG-1999; 98US-00380142.
PR 29-OCT-1999; 98US-0162506P.
PR 30-NOV-1999; 98WO-US028313.
PR 02-DEC-1999; 98WO-US028551.
PR 02-DEC-1999; 98WO-US028565.
PR 16-DEC-1999; 98WO-US030935.
PR 30-DEC-1999; 98WO-US031243.
PR 30-DEC-1999; 98WO-US031274.
PR 05-JAN-2000; 2000WO-US0001219.
PR 06-JAN-2000; 2000WO-US000277.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.

PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000US-00709238.
PR 27-NOV-2000; 2000US-00723749.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00815744.
PR 22-MAR-2001; 2001US-00816920.
PR 22-MAR-2001; 2001WO-US009552.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 21-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.
PR 19-JUN-2001; 2001US-00886342.
PR 20-JUN-2001; 2001WO-US019692.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-JUL-2001; 2001WO-US021735.
PR 30-JUL-2001; 2001US-00918585.
XX
XX (GETH) GENENTECH INC.
PA
XX Ashkenazi AJ, Baker KP, Botstein D, Deencyers L, Eaton DJ,
XX Best Local Similarity 100.0%; Pred. No. 0;
PI Ferrara N, Filvarcoff E, Fong S, Gao W, Gerber H, Gerlitsen ME;

Query Match 100.0%; Score 3843; DB 6; Length 730;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCORTRAPMKKESQTERAALGFRRGSGMFASGNNQVPIEAGSMALLPLPLLL 60
DB 1 MGVCORTRAPMKKESQTERAALGFRRGSGMFASGNNQVPIEAGSMALLPLPLLL 60
QY PLILLIKLHMPQIRKMLPADIAFAVRALCKRALRARALAAAADPEPGGCSLARLAE 120
DB PLILLIKLHMPQIRKMLPADIAFAVRALCKRALRARALAAAADPEPGGCSLARLAE 120
QY 61 PLILLIKLHMPQIRKMLPADIAFAVRALCKRALRARALAAAADPEPGGCSLARLAE 120
DB 61 PLILLIKLHMPQIRKMLPADIAFAVRALCKRALRARALAAAADPEPGGCSLARLAE 120
QY 121 LAQORAAHTFLIHGSRFSTSEARESRNRAARFALAMDMDPDGDSGSGAGCERA 180
DB 121 LAQORAAHTFLIHGSRFSTSEARESRNRAARFALAMDMDPDGDSGSGAGCERA 180
QY 181 APGAGDAAAGSAGAEFAGDGAARGGGAAAPLSGATVALLPAPGPFILWIFGLAVAGR 240
DB 181 APGAGDAAAGSAGAEFAGDGAARGGGAAAPLSGATVALLPAPGPFILWIFGLAVAGR 240
QY 241 TAFVPTALRRGPLHLHLCRSCGAAVLAPLEFLESPDIPALRAWGLHMAAGPGTHPG 300
DB 241 TAFVPTALRRGPLHLHLCRSCGAAVLAPLEFLESPDIPALRAWGLHMAAGPGTHPG 300
QY 301 ISDLAEVSAEVDGAPPGYLSPOSITDTCIVIFTSGLTGLPFAARISHKLILQCCGFYQ 360
DB 301 ISDLAEVSAEVDGAPPGYLSPOSITDTCIVIFTSGLTGLPFAARISHKLILQCCGFYQ 360
QY 361 LCGVHOEDVYIYALPLYMNSGLIGVCMGIGATVYLSKSSAGAFMEDCCQHRVTYQ 420
DB 361 LCGVHOEDVYIYALPLYMNSGLIGVCMGIGATVYLSKSSAGAFMEDCCQHRVTYQ 420
QY 421 YIGELCRYVNOPPSAERGHKVRILAVSGLRPDTERRVRRRGPQVLETYGLTEGNYA 480
DB 421 YIGELCRYVNOPPSAERGHKVRILAVSGLRPDTERRVRRRGPQVLETYGLTEGNYA 480
QY 481 TINYTGQRGAVGASWLYKHIPPSLIRYDVTTGEPDIRDPQHCVAATSGEPGLVAPVS 540
DB 481 TINYTGQRGAVGASWLYKHIPPSLIRYDVTTGEPDIRDPQHCVAATSGEPGLVAPVS 540
QY 541 QOSPFLGYAGGPELAAQSKLLKDFRRGVDVFPNTGLLVCDGQFLRFHRTGTPFWKGB 600
DB 541 QOSPFLGYAGGPELAAQSKLLKDFRRGVDVFPNTGLLVCDGQFLRFHRTGTPFWKGB 600

Db 541 QOSPFLGVAGPELAQKLLKDVFRPGDVFFNTGDLVCDQGFLEFRHDTGTFRWKE 600
Qy NATTEVEVFEFALDPLFLOEVNVYGVTVPGHESRAMALVVRPPALDLMQLYTHVSEN 660
Db 601 NATTEVEVFEFALDPLFLOEVNVYGVTVPGHESRAMALVVRPPALDLMQLYTHVSEN 660
Qy 661 PPARPRFRLQESLATTETFPKQKVMANEGFDPSTLSDPLVLDQVAGVPLPTARY 720
Db 661 PPARPRFRLQESLATTETFPKQKVMANEGFDPSTLSDPLVLDQVAGVPLPTARY 720
Qy 721 SALLAGNLR 730
Db 721 SALLAGNLR 730
RESULT 15
ID ABO19547 standard; protein; 730 AA.
AC ABO19547;
XX
XX 27-AUG-2003 (first entry)
DE Novel human secreted and transmembrane polypeptide #15.
XX
XX Human; secreted and transmembrane protein; PRO; viral infection;
KW tumour growth; retinal disorder; injury; sight loss;
KW retinitis pigmentosa; age-related macular degeneration;
KW sport-related joint problem; articular cartilage defect; osteoarthritis;
KW rheumatoid arthritis; wound healing; obesity; diabetes; insulinemia;
KW kidney disorder; mesangial cell function; Berger disease; nephropathy;
KW cardiac disease; dermatitis; Crohn disease; neuropathy;
KW diabetic peripheral neuropathy; autonomic neuropathy;
KW reduced motility of the gastrointestinal tract;
KW Charcot-Marie-Tooth disease; Fabry's disease; Tangier disease;
KW Refsum's disease.
XX
XX Homo sapiens.
OS
XX
XX US2003049633-A1.
PD 13-MAR-2003.
XX
PF 16-OCT-2001; 2001US-00978585.
XX
XX 17-OCT-1997; 97US-0062250P.
PR 03-NOV-1997; 97US-0064249P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066364P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077641P.
PR 11-MAR-1998; 98US-0077649P.
PR 12-MAR-1998; 98US-0077991P.
PR 13-MAR-1998; 98US-0078004P.
PR 17-MAR-1998; 98US-0084022P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078910P.
PR 20-MAR-1998; 98US-0078936P.
PR 20-MAR-1998; 98US-0078939P.
PR 25-MAR-1998; 98US-0079294P.
PR 26-MAR-1998; 98US-0079656P.
PR 27-MAR-1998; 98US-0079663P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079689P.
PR 27-MAR-1998; 98US-0079728P.
PR 27-MAR-1998; 98US-0079786P.
PR 30-MAR-1998; 98US-0079920P.
PR 30-MAR-1998; 98US-0079923P.
PR 31-MAR-1998; 98US-0080105P.
PR 31-MAR-1998; 98US-0080107P.

PR 31-MAR-1998; 98US-0080165P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080328P.
PR 01-APR-1998; 98US-0080332P.
PR 01-APR-1998; 98US-0080334P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 08-APR-1998; 98US-0081071P.
PR 09-APR-1998; 98US-0081195P.
PR 09-APR-1998; 98US-0081203P.
PR 09-APR-1998; 98US-0081229P.
PR 15-APR-1998; 98US-0081817P.
PR 15-APR-1998; 98US-0081819P.
PR 15-APR-1998; 98US-0081838P.
PR 15-APR-1998; 98US-0081952P.
PR 15-APR-1998; 98US-0081955P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082700P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 22-APR-1998; 98US-0082804P.
PR 23-APR-1998; 98US-0082796P.
PR 27-APR-1998; 98US-0083336P.
PR 28-APR-1998; 98US-0083322P.
PR 28-APR-1998; 98US-0083392P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083500P.
PR 29-APR-1998; 98US-0083545P.
PR 29-APR-1998; 98US-0083554P.
PR 29-APR-1998; 98US-0083558P.
PR 29-APR-1998; 98US-0083559P.
PR 30-APR-1998; 98US-0083742P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 06-MAY-1998; 98US-0084441P.
PR 07-MAY-1998; 98US-0084598P.
PR 07-MAY-1998; 98US-0084600P.
PR 07-MAY-1998; 98US-0084627P.
PR 07-MAY-1998; 98US-0084637P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 13-MAY-1998; 98US-0085323P.
PR 13-MAY-1998; 98US-0085338P.
PR 13-MAY-1998; 98US-0085339P.
PR 15-MAY-1998; 98US-0085573P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085800P.
PR 15-MAY-1998; 98US-0085882P.
PR 15-MAY-1998; 98US-0085897P.
PR 15-MAY-1998; 98US-0085970P.
PR 15-MAY-1998; 98US-0085704P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086414P.
PR 22-MAY-1998; 98US-0086430P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087106P.
PR 28-MAY-1998; 98US-0087208P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 30-JUL-1998; 98US-0091359P.
PR 30-JUL-1998; 98US-0094651P.
PR 11-SEP-1998; 98US-0100038P.
PR 07-OCT-1998; 98US-00168978.
PR 07-OCT-1998; 98WO-US021141.

PR 02-NOV-1998; 98US-00184216.
PR 06-NOV-1998; 98US-00187368.
PR 20-NOV-1998; 98US-0109304P.
PR 20-NOV-1998; 98WO-US024855.
PR 20-DEC-1998; 98US-00202054.
PR 22-DEC-1998; 98US-00218517.
PR 22-DEC-1998; 98US-0113296P.
PR 23-DEC-1998; 98US-0113621P.
PR 05-JAN-1999; 99WO-US000106.
PR 05-MAR-1999; 99WO-US025465.
PR 08-MAR-1999; 99WO-US005028.
PR 10-MAR-1999; 99WO-US005686.
PR 12-MAR-1999; 99WO-US005190.
PR 12-MAR-1999; 99US-00267213.
PR 12-MAR-1999; 99US-0123757P.
PR 29-MAR-1999; 99US-0126773P.
PR 12-APR-1999; 99US-00284231.
PR 21-APR-1999; 99US-0130232P.
PR 26-APR-1999; 99US-0131022P.
PR 28-APR-1999; 99US-0131445P.
PR 14-MAY-1999; 99US-00311832.
PR 14-MAY-1999; 99US-0134287P.
PR 14-MAY-1999; 99WO-US010733.
PR 02-JUN-1999; 99WO-US012252.
PR 16-JUN-1999; 99US-0139557P.
PR 23-JUN-1999; 99US-0141037P.
PR 07-JUL-1999; 99US-0142680P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 25-AUG-1999; 99US-00380137.
PR 25-AUG-1999; 99US-00380138.
PR 25-AUG-1999; 99US-00380142.
PR 29-OCT-1999; 99US-0162506P.
PR 30-NOV-1999; 99WO-US028313.
PR 02-DEC-1999; 99WO-US028551.
PR 02-DEC-1999; 99WO-US028565.
PR 16-DEC-1999; 99WO-US030095.
PR 30-DEC-1999; 99WO-US031243.
PR 30-DEC-1999; 99WO-US031274.
PR 05-JAN-2000; 2000WO-US000219.
PR 06-JAN-2000; 2000WO-US000277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US003565.
PR 18-FEB-2000; 2000WO-US004341.
PR 24-FEB-2000; 2000WO-US005004.
PR 02-MAR-2000; 2000WO-US005841.
PR 10-MAR-2000; 2000WO-US006319.
PR 21-MAR-2000; 2000WO-US007532.
PR 30-MAR-2000; 2000WO-US008439.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015254.
PR 28-JUL-2000; 2000WO-US020710.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000US-00792328.
PR 27-NOV-2000; 2000US-00723749.
PR 01-DEC-2000; 2000WO-US02678.
PR 20-DEC-2000; 2000US-00747259.
PR 28-FEB-2001; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 22-MAR-2001; 2001US-00816744.
PR 22-MAR-2001; 2001US-00816920.
PR 22-MAR-2001; 2001WO-US009552.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001WO-US017092.
PR 01-JUN-2001; 2001US-00872035.
PR 01-JUN-2001; 2001WO-US017800.
PR 05-JUN-2001; 2001US-00874503.
PR 14-JUN-2001; 2001US-00882636.

Query Match 100.0%; Score 3843; DB 6; Length 730;

	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches	730;	Conservative	0;	Indels	0;	Gaps	0;
QY	1	MGVCORTRAPWKEKSLTERAALGFRKSGSGMFASGNNQVPIIEBAGSMALLPLILL	60								
Db	1	MGVCORTRAPWKEKSLTERAALGFRKSGSGMFASGNNQVPIIEBAGSMALLPLILL	60								
QY	61	PLILLKLHMPQIRMLPADLAFVRALCCRRALRARAALAAAADPEPGGCSLANRLAE	120								
Db	61	PLILLKLHMPQIRMLPADLAFVRALCCRRALRARAALAAAADPEPGGCSLANRLAE	120								
QY	121	LAQCPAAHTFLIHGSRFSYSSEAEBSNBARAFRALGMDGPDGDSGSGSAGEGERA	180								
Db	121	LAQCPAAHTFLIHGSRFSYSSEAEBSNBARAFRALGMDGPDGDSGSGSAGEGERA	180								
QY	181	APGAGDAAAGSGAEFGAGDGAARGGGAAPLSPGATVALLPAGPEFLMFGLAAGAR	240								
Db	181	APGAGDAAAGSGAEFGAGDGAARGGGAAPLSPGATVALLPAGPEFLMFGLAAGAR	240								
QY	241	TAFVPTALRRGPLHLHCLBSCGARALVLAPEFLESLEPDLPAIRAMGLHMAAGPTHPAG	300								
Db	241	TAFVPTALRRGPLHLHCLBSCGARALVLAPEFLESLEPDLPAIRAMGLHMAAGPTHPAG	300								
QY	301	ISDLAEVSAEVDGFPYGLSSPOSITPDCLYIFTSGTGLPKAARISHKLIOCGFPQ	360								
Db	301	ISDLAEVSAEVDGFPYGLSSPOSITPDCLYIFTSGTGLPKAARISHKLIOCGFPQ	360								
QY	361	LCGVHEDVIYIALPLYHNGSILGIVCMGIGATVVLKSKTSAGOFMEDCOOHRVTYQ	420								
Db	361	LCGVHEDVIYIALPLYHNGSILGIVCMGIGATVVLKSKTSAGOFMEDCOOHRVTYQ	420								
QY	421	YIGELCRYLVNOPSKAERGHKVRILAVGSLRDPDTERFVRBFGPIQVJETYGLTGNTVA	480								
Db	421	YIGELCRYLVNOPSKAERGHKVRILAVGSLRDPDTERFVRBFGPIQVJETYGLTGNTVA	480								
QY	481	TINYTGORGAVGRASWLYXHIFFPSLIRYDVTGEBIRDPQGHCMATSGEPGLVAPVS	540								
Db	481	TINYTGORGAVGRASWLYXHIFFPSLIRYDVTGEBIRDPQGHCMATSGEPGLVAPVS	540								
QY	541	QOSPFLGYAGGPELLAAGKLLKDYFRPGDVFPNTGDLVCDGDFLRPHDRGTDPFWKGE	600								
Db	541	QOSPFLGYAGGPELLAAGKLLKDYFRPGDVFPNTGDLVCDGDFLRPHDRGTDPFWKGE	600								
QY	601	NVATTEVAEVEFALDLOEVNNGVTVPGHEGRAGAAALVLRPHALDLMQYTHVSENL	660								
Db	601	NVATTEVAEVEFALDLOEVNNGVTVPGHEGRAGAAALVLRPHALDLMQYTHVSENL	660								
QY	661	PYARPPFLRLQSLATTETPKQCKVRANEGDPSTLSDPLVLDQAVGAYLPLTTARY	720								
Db	661	PYARPPFLRLQSLATTETPKQCKVRANEGDPSTLSDPLVLDQAVGAYLPLTTARY	720								
QY	721	SALLAGNLR 730									
Db	721	SALLAGNLR 730									

Search completed: June 30, 2004, 17:30:37
Job time : 65 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2004, 17:30:45 ; Search time 52 Seconds

(without alignments)
4429.390 Million cell updates/sec

Title: US-10-030-226-2

Sequence: 1 MGVCGRTPWKEKESQLERA.....AYLPITARYSALLAGLRLI 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriopl:*
- 17: sp_archaeopl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3843	100.0	730	4	Q9BTY5
2	3683	95.8	700	4	Q9BTY5 homo sapien
3	2763	71.9	730	4	Q8N2X7
4	2103	54.7	446	11	Q88561
5	2083.5	54.2	402	4	Q8BTU5
6	1407	36.6	288	4	Q8TBU0
7	1351.5	35.2	619	4	Q7Z6E6
8	1350.5	35.1	619	4	Q86YF6
9	1346	35.0	620	11	Q9Y2P4
10	1346	34.2	625	13	Q8AVC5
11	1312.5	33.9	690	4	Q9Y2P5
12	1303.5	32.6	662	11	Q88694
13	1251	32.6	689	11	Q91VD5
14	1231	32.0	690	11	Q9E838
15	1193.5	31.1	623	11	Q88560

17	1076	28.0	638	16	Q93GR0	Q93GR0 bradyrhizob
18	1048.5	27.3	643	11	Q91VE0	Q91VE0 mus musculu
19	1043	27.1	604	16	Q89CL2	Q89CL2 bradyrhizob
20	1028	26.7	635	16	Q9A5Z8	Q9A5Z8 caulobacter
21	995	25.9	687	5	Q85XR7	Q85XR7 drosophila
22	992.5	25.8	661	5	Q9W185	Q9W185 drosophila
23	985.5	25.6	506	11	Q88562	Q88562 mus musculu
24	952	24.8	626	5	Q9YKX1	Q9YKX1 drosophila
25	942.5	24.5	641	4	Q95186	Q95186 homo sapien
26	930.5	24.2	679	5	Q9W1W0	Q9W1W0 drosophila
27	924.5	24.1	679	5	Q811F9	Q811F9 drosophila
28	903	23.5	608	16	Q8HZV4	Q8HZV4 pseudomonas
29	859.5	22.4	654	5	Q85XT7	Q85XT7 drosophila
30	858	22.3	465	5	Q18916	Q18916 caenorhabdi
31	841.5	21.9	597	16	Q05307	Q05307 mycobacteri
32	841.5	21.9	597	16	Q7U0E2	Q7U0E2 mycobacteri
33	839.5	21.8	650	5	Q18878	Q18878 caenorhabdi
34	763.5	19.9	334	11	Q9CV67	Q9CV67 mus musculu
35	725.5	18.9	643	3	Q42633	Q42633 cochlioniu
36	711	18.5	609	3	Q8JOE9	Q8JOE9 cephalospor
37	669	17.4	669	3	Q6D021	Q6D021 saccharomyc
38	540.5	14.1	258	4	Q9BTX1	Q9BTX1 homo sapien
39	469.5	12.2	561	16	Q7WPV1	Q7WPV1 bordetella
40	466.5	12.1	561	16	Q7WBV5	Q7WBV5 bordetella
41	453	11.8	554	16	Q89PP7	Q89PP7 bradyrhizob
42	411	10.7	559	16	Q7W037	Q7W037 bordetella
43	406.5	10.6	522	16	Q83MG9	Q83MG9 shigella fl
44	400	10.5	502	16	Q7WPM7	Q7WPM7 bordetella
45	400	10.4	502	16	Q7WBN8	Q7WBN8 bordetella

ALIGNMENTS

RESULT 1

Q9BTY5 PRELIMINARY; PRT; 730 AA.

AC Q9BTY5; Q96SM5;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Hypothetical protein FLJ14593 (Hypothetical protein
 NT2RP200142).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 31-730 FROM N.A.
 RA Isogai T., Ota T., Hayaishi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NED human cDNA sequencing project";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
 RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,
 RA Nagahari K., Sugano S., Isogai T.;
 RT "HRI human cDNA sequencing project";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC003041; AA03041.1; -
 DR EMBL; BC009916; AA09916.1; -
 DR EMBL; AK027499; BAB5156.1; -

DR EMBL; AK075377; BAC11578.1; --
 DR Genew; HGNC:10997; SLCT2A3.
 DR GO; GO:0003824; F: catalytic activity; IEA.
 DR GO; GO:0008152; P: metabolism; IEA.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 2.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 DR Hypothetical protein.
 KW SEQUENCE 730 AA; 78644 MW; E8CC4E9463A534A0 CRC64;

Query Match 100.0%; Score 3843; DB 4; Length 730;
 Best Local Similarity 100.0%; Pred. No. 1,1e-255;
 Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NCVCORTRAPEKESQLEBALGFRKSGMFASSMNOVTVEEAGSMAALLPLILL 60
 DB 1 NCVCORTRAPEKESQLEBALGFRKSGMFASSMNOVTVEEAGSMAALLPLILL 60
 QY 61 PLILLKHLMPOLRMLPADLAFVAVLCCCKALRALAAAAADPEGEGCSLAMP 120
 DB 61 PLILLKHLMPOLRMLPADLAFVAVLCCCKALRALAAAAADPEGEGCSLAMP 120
 QY 121 LAQQAATFTLHGSRRESYSSEARESNRAAFALGMDWGPDDGSGSAGEG 180
 DB 121 LAQQAATFTLHGSRRESYSSEARESNRAAFALGMDWGPDDGSGSAGEG 180
 QY 181 AAGADDAAGSAGEFNGDGAAGGGAAPISPGATVALLPAGPEFLMFGAKGLR 240
 DB 181 AAGADDAAGSAGEFNGDGAAGGGAAPISPGATVALLPAGPEFLMFGAKGLR 240
 QY 241 TAFVFTALRRGPLLHCLSCGALVLAPEFLESLEPDLPALRAMGLHMAAGGT 300
 DB 241 TAFVFTALRRGPLLHCLSCGALVLAPEFLESLEPDLPALRAMGLHMAAGGT 300
 QY 301 ISDILAESAEDVGPVGLSSPGSITDTCLYTFSSGTGLPKARISHLILCOQGF 360
 DB 301 ISDILAESAEDVGPVGLSSPGSITDTCLYTFSSGTGLPKARISHLILCOQGF 360
 QY 361 LCGVHOEDVLYALPLVHNSGSLGIVGCMGIGATVILKSKFSAGQFEDCOHRVTF 420
 DB 361 LCGVHOEDVLYALPLVHNSGSLGIVGCMGIGATVILKSKFSAGQFEDCOHRVTF 420
 QY 421 YIGELCRYLVNPPSKAERGHKVRILAVSGSLPDTWEPVRRFGPLQYLETG 480
 DB 421 YIGELCRYLVNPPSKAERGHKVRILAVSGSLPDTWEPVRRFGPLQYLETG 480
 QY 481 TINYTGORGAVRASMLYKHIFPSLIRYDTTGPIDPOGHCMATSPGEGILVAV 540
 DB 481 TINYTGORGAVRASMLYKHIFPSLIRYDTTGPIDPOGHCMATSPGEGILVAV 540
 QY 541 QOSPFLGYAGBELAQGLKLDVFRPGDVFPNTGDLVCDGQFLRFHRTGDTFR 600
 DB 541 QOSPFLGYAGBELAQGLKLDVFRPGDVFPNTGDLVCDGQFLRFHRTGDTFR 600
 QY 601 NVAITEVAVFEPALFLQEVNYYGVTFVGHGRAMALVLRPPALDLMOLYHVS 660
 DB 601 NVAITEVAVFEPALFLQEVNYYGVTFVGHGRAMALVLRPPALDLMOLYHVS 660
 QY 661 PVPARPRFLRLOESLATTETFKQKVRMANEGFDSSTLDPYLVDQAVGAYPL 720
 DB 661 PVPARPRFLRLOESLATTETFKQKVRMANEGFDSSTLDPYLVDQAVGAYPL 720
 QY 721 SALLAGNLRI 730
 DB 721 SALLAGNLRI 730

RESULT 2

ID Q8N2X7 PRELIMINARY; PRT: 700 AA.
 AC Q8N2X7; 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DR Hypothetical protein.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NX NCBI_Taxid=9606;
 RN [1]
 RE SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (May-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; BC029792; AAH29792.1; --
 DR GO; GO:0003824; F: catalytic activity; IEA.
 DR GO; GO:0008152; P: metabolism; IEA.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 2.
 DR PROSITE; PS00455; AMP_BINDING; 1.
 DR Hypothetical protein.
 KW SEQUENCE 700 AA; 75356 MW; 4BE800A1D872A51A CRC64;

Query Match 95.8%; Score 3683; DB 4; Length 700;
 Best Local Similarity 100.0%; Pred. No. 1e-244;
 Matches 700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 MEASGNNQVPIEEAGSMAALLPLILLPLILLKHLMPOLRMLPADLAFVAVLCC 90
 DB 1 MEASGNNQVPIEEAGSMAALLPLILLPLILLKHLMPOLRMLPADLAFVAVLCC 90
 QY 91 RAIRALAAAAADPEGEGCSLAMPRLAEALQQAATFTLHGSRRESYSARESNRA 150
 DB 91 RAIRALAAAAADPEGEGCSLAMPRLAEALQQAATFTLHGSRRESYSARESNRA 150
 QY 151 ARAFLALGMDWGPDDGSGSAGEGERAPAGDAAAGSAEFGAGDGAAGGAAAP 210
 DB 151 ARAFLALGMDWGPDDGSGSAGEGERAPAGDAAAGSAEFGAGDGAAGGAAAP 210
 QY 211 LSPGATVALLPAGPEFLMFGAKGLRRTAFVFTALRRGPLLHCLSCGALVLAPE 270
 DB 211 LSPGATVALLPAGPEFLMFGAKGLRRTAFVFTALRRGPLLHCLSCGALVLAPE 270
 QY 271 FLESLEPDLPALRAMGLHMAAGGTTHPAGISDILAESAEDVGPVGLSSPG 330
 DB 271 FLESLEPDLPALRAMGLHMAAGGTTHPAGISDILAESAEDVGPVGLSSPG 330
 QY 331 LYFTSGTGLPKARISHLILCOQGFYQLCGVHOEDVLYALPLVHNSGSLGIVGM 390
 DB 331 LYFTSGTGLPKARISHLILCOQGFYQLCGVHOEDVLYALPLVHNSGSLGIVGM 390
 QY 391 GIGATVILKSKFSAGQFEDCOHRVTFQYIGELCRYLVNPPSKAERGHKVRILAV 450
 DB 391 GIGATVILKSKFSAGQFEDCOHRVTFQYIGELCRYLVNPPSKAERGHKVRILAV 450
 QY 451 LREDTWERFVRFPGLQVLETYGLTEGNVATINYTGORGAVRASMLYKHIFPSLIR 510
 DB 451 LREDTWERFVRFPGLQVLETYGLTEGNVATINYTGORGAVRASMLYKHIFPSLIR 510
 QY 511 YTGGERIRPOGHCMATSPGEGILVAVYSOOSPFLGYAGBELAQGLKLDVFRPG 570
 DB 511 YTGGERIRPOGHCMATSPGEGILVAVYSOOSPFLGYAGBELAQGLKLDVFRPG 570
 QY 571 FNTGDLVCDGQFLRFHRTGDTFRKGENVAITEVAEFAALDFLOEVNYYGVTFV 630
 DB 571 FNTGDLVCDGQFLRFHRTGDTFRKGENVAITEVAEFAALDFLOEVNYYGVTFV 630
 QY 631 EGRAGMAALVLRPPALDLMOLYHVS ENULPYAPRPRFLRLOESLATTETFKQKVR 690
 DB 631 EGRAGMAALVLRPPALDLMOLYHVS ENULPYAPRPRFLRLOESLATTETFKQKVR 690
 QY 691 EGFDPSTLSDPLYLVDQAVGAYPLPTTARYSALLAGNLRI 730
 DB 691 EGFDPSTLSDPLYLVDQAVGAYPLPTTARYSALLAGNLRI 730


```

RESULT 3
088561 PRELIMINARY; PRT; 614 AA.
AC 088561;
DR 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Fatty acid transport protein 3 (FATP3) (long-chain fatty acid
  transport protein 3) (Fragment).
GN SLC27A3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=98337965; PubMed=9671728;
RA Hirsch D., Stahl A., Lodish H.F.;
RT "A family of fatty acid transporters conserved from mycobacterium to
  man."
RL Proc. Natl. Acad. Sci. U.S.A. 95:8625-8629(1998).
CC -|- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS
  ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING
  AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES
  IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR
  TRIGLYCERIDE SYNTHESIS.
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PLASMA MEMBRANE.
CC -|- TISSUE SPECIFICITY: LUNG, LIVER, AND TESTIS.
CC -|- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT
  COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
EMBL AF072758; AAC40187.1; -.
MD MGD; MGI:1347358; SLC27A3.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0006869; P:lipid transport; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 2.
DR PROSITE: PS00455; AMP BINDING; 1.
KM Glycoprotein; lipid transport; Transmembrane; Transport.
FT NON TER 1 1
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 262 282 POTENTIAL.
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 614 AA; 67041 MW; 33C2A558CDD9D989 CRC64;

Query Match 71.9%; Score 2763; DB 11; Length 614;
Best Local Similarity 83.5%; Pred. No. 1,5e-181;
Matches 526; Conservative 35; Mismatches 53; Indels 16; Gaps 2;

QY 101 AAADPEGEGGCSLAWRLAEIAQCPAAFTFLIHGSRFSYSEAEERNSRARAFRLALGW 160
DB 1 AAADPESSSGCSLAWRLAYLAREQPTHFLIHGQRFSYAEAEESRIRARAFRLARAGW 60
QY 161 DMGPFGSGSGBSAEGGERAAPGADAAAGSAEPAGGDGAARGGGAAPSPGATVAL 220
DB 61 TGGRRG--SGRGSTEGRVAPAPGDA--RRTTAPPLAPGATVAL 104
QY 221 LPAGEFELMWFGKAGKRTAFVPTALRPRGLHCLRS CGARLVLAPEFLSELEPPLP 280
DB 105 LPAGDFELMWFGKAGKRTAFVPTALRPRGLHCLRS CGASALVLAPEFLSELEPPLP 164
QY 281 ALRANGHLWMAAGPETHPAGISDLIAEVAEYDGVPGYLS POSITDTCLYFTSGTTG 340
DB 165 ALRANGHLWMAAGPETHPAGISDLIAEVAEYDGVPGYLS POSIMDTCLYFTSGTTG 224
QY 341 LFKARISHLKLQCGFYQCGVHQBVDYLYALPLYMGSGLGIYVCMGIGATVVLKS 400
DB 225 LFKARISHLKLQCGFYHLCGHQBVDYLYALPLYMGSGLGIYVGLGATVVLKP 284
QY 401 KFSAGQFEDCCQGHVTVFOYIGELCRYLVNPPSKAERGHKVRILAVSGSLRPDTERFV 460
DB 285 KFSAGQFMDCCQKHVTVFOYIGELCRYLVNPPSKAEFDHKVRLAVSGSLRPDTERFL 344

```

```

RESULT 4
08BK70 PRELIMINARY; PRT; 446 AA.
AC 08BK70;
DR 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Solute carrier family 27.
GN SLC27A3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=body;
RX MEDLINE=42354683; PubMed=12466851;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RA "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK076014; BAC36120.1; -.
DR MGD; MGI:1347358; SLC27A3.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP BINDING; 1.
SQ SEQUENCE 446 AA; 49317 MW; BAIED75849EDF92B CRC64;

Query Match 54.7%; Score 2103; DB 11; Length 446;
Best Local Similarity 87.9%; Pred. No. 2,3e-136;
Matches 392; Conservative 26; Mismatches 28; Indels 0; Gaps 0;

QY 285 MGLHMAAGPETHPAGISDLIAEVAEYDGVPGYLS POSITDTCLYFTSGTTGLPFA 344
DB 1 MGLHMAAGPETHPAGISDLIAEVAEYDGVPGYLSAQNIMDTCLYFTSGTTGLPFA 60
QY 345 ARISHLKLQCGFYQCGVHQBVDYLYALPLYMGSGLGIYVCMGIGATVVLKSFSA 404
DB 61 ARISHLKLQCGFYHLCGHQBVDYLYALPLYMGSGLGIYVGLGATVVLKSFSA 120
QY 405 GQFMECCQGHVTVFOYIGELCRYLVNPPSKAERGHKVRILAVSGSLRPDTERFVRFG 464
DB 121 GQFMECCQKHVTVFOYIGELCRYLVNPPSKAECHKVRILAVSGSLRPDTERFLRRG 180
QY 465 PLQVLETYGTEGNVATINTYTGQGA VGRASWLKYHIFPFSILRYDVTGEPIDPQGHG 524
DB 181 PLQVLETYGTEGNVATINTYTGQGA VGRASWLKYHIFPFSILRYDVTGEPIDPQGHG 240

```

```

QY 525 MATSPREPLVAPVSOQSPFLGYAGPELAQKLLKDVFRPGDVFNNTGDLVCDQGF 584
DB 241 MTTSPEEPPLVAPVSOQSPFLGYAGPELAQKLLKDVFRPGDVFNNTGDLVCDQGF 300
QY 585 LRFHRTGDTFRMKGENVATTEVAEVEFALDPLQEVNYYGVTPGHEGRAGMAALVLRP 644
DB 301 LRFHRTGDTFRMKGENVATTEVAEVEFALDPLQEVNYYGVTPGHEGRAGMAALVLRP 360
QY 645 HALDLMLQYTHYSENPYPARPRFLRLOESLATTTFYQOKVRMANEGFDPSITLSDPLY 704
DB 361 QALNTLVQYSHVSENPYPARPRFLRLOESLATTTFYQOKVRMANEGFDPSITLSDPLY 420
QY 705 LDOAVGAYPLFTFARYSALLAGNLRI 730
DB 421 LDOIGAYPLFTFARYSALLAGNLRI 446

RESULT 5
Q9BTJ5 PRELIMINARY; PRT; 402 AA.
ID Q9BTJ5;
AC Q9BTJ5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to hypothetical protein WC4365 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003654; AA03654.1;
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP BINDING; 1.
KW Hypothetical protein.
FT NON TER
SQ SEQUENCE 402 AA; 44092 MW; B1D946FC9822C3F7 CRC64;

Query Match 54.2%; Score 2083.5; DB 4; Length 402;
Best Local Similarity 91.7%; Pred. No. 4, 4e-135;
Matches 398; Conservative 1; Mismatches 0; Indels 35; Gaps 1;

QY 297 HPAGISDLAESAEDVGPVPGVLSPOSITDTCTCYIFTSCTTGPKAARISHLKILQCO 356
DB 4 HPAGISDLAESAEDVGPVPGVLSPOSITDTCTCYIFTSCTTGPKAARISHLKILQCO 63
QY 357 GFYQLGQHEDEVLYLALPLYMSSLLGIVGCMGIGATVVKSFSAQGFVEDCQQRHV 416
DB 64 GFYQLGQHEDEVLYLALPLYMSSLLGIVGCMGIGATVVKSFSAQGFVEDCQQRHV 123
QY 417 TYFQYIGELCRYLVNPPSKARHGKVRILAVSGSLRPDTEWRFVRFGPIQVLETYGLTE 476
DB 124 TYFQYIGELCRYLVNPPSKARHGKVRILAVSGSLRPDTEWRFVRFGPIQVLETYGLTE 183
QY 477 GNVATINNYGORGAVGRASWLYKHIFPESLIYDYVTGEPRIIDPOGHOMATSPGEGLLV 536
DB 184 GNVATINNYGORGAVGRASWLYKHIFPESLIYDYVTGEPRIIDPOGHOMATSPGEGLLV 243
QY 537 APVSQSQSPFLGYAGPELAQKLLKDVFRPGDVFNNTGDLVCDQGFRLRFHRTGDTFR 596
DB 244 APVSQSQSPFLGYAGPELAQKLLKDVFRPGDVFNNTGDLVCDQGFRLRFHRTGDTFR 303
QY 597 MKGENVATTEVAEVEFALDPLQEVNYYGVTPGHEGRAGMAALVLRPHALDMLQYTHV 656
DB 304 MKGENVATTEVAEVEFALDPLQEVNYYGVTPGHEGRAGMAALVLRPHALDMLQYTHV 363
QY 657 SENLPPYARPRFLRLOESLATTTFYQOKVRMANEGFDPSITLSDPLYLDOAVGAYPLFT 716

```

```

DB 364 SENLPPYARPRFLRLOESLATTTFYQOKVRMANEGFDPSITLSDPLYLDOAVGAYPLFT 388
QY 717 TARYSALLAGNLRI 730
DB 389 TARYSALLAGNLRI 402

RESULT 6
Q9TEJ0 PRELIMINARY; PRT; 288 AA.
ID Q9TEJ0;
AC Q9TEJ0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE FLJ00207 protein (fragment).
GN FLJ00207.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Ikuya H., Takano U., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human spleen."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074134; BAB84960.1;
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 2.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
FT NON TER
SQ SEQUENCE 288 AA; 29865 MW; 09AD767A6497A95C CRC64;

Query Match 36.6%; Score 1407; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 9e-89;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVCQTRAPWKEKSLERALGFRKSGSGMFASGNNQVPIEASGNALLLPLLLI 60
DB 13 MGVCQTRAPWKEKSLERALGFRKSGSGMFASGNNQVPIEASGNALLLPLLLI 72
QY 61 PLILKILHLPQILRMLPADIAFVYRALCKRAIRALAAAAADPEPGGGLARLAE 120
DB 73 PLILKILHLPQILRMLPADIAFVYRALCKRAIRALAAAAADPEPGGGLARLAE 132
QY 121 LAQORAAHTFLHSGRRFSYSEARERSENRAAFLEALGMDWGPDDGSGSGAGEGERA 180
DB 133 LAQORAAHTFLHSGRRFSYSEARERSENRAAFLEALGMDWGPDDGSGSGAGEGERA 192
QY 181 APGAGDAAGSGAEFAGGDAAGGAAAPLSGATVALLPAGPEFLWLFGLAKGLR 240
DB 193 APGAGDAAGSGAEFAGGDAAGGAAAPLSGATVALLPAGPEFLWLFGLAKGLR 252
QY 241 TAFVPTALRRGRLHCLRSGARALVYAP 269
DB 253 TAFVPTALRRGRLHCLRSGARALVYAP 281

RESULT 7
Q7Z6E6 PRELIMINARY; PRT; 619 AA.
ID Q7Z6E6;
AC Q7Z6E6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKFZP79M0564.
GN DKFZP79M0564.
OS Homo sapiens (Human).

```


OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Steinberg S.J., Watkins P.A.;
 RT "Human Very Long-Chain Acyl-CoA Synthetase Homolog 1."
 RL Submitted (Mar-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF064254; AAD29443.1;
 DR GO; GO:0004467; P:long-chain-fatty-acid-CoA ligase activity; TAS.
 DR GO; GO:000038; P:very-long-chain fatty acid metabolism; TAS.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF0501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP-BINDING; 1.
 DR SEQUENCE 619 AA; 7011 MW; 20264CE3FB44FF5 CRC64;
 SQ
 Query Match 35.1%; Score 1350.5; DB 4; Length 619;
 Best Local Similarity 43.8%; Pred. No. 1.9e-84;
 Matches 267; Conservative 103; Mismatches 186; Indels 53; Gaps 3;
 QY 122 ACGRAAHTLIGSRFSESEAESNRAPAFALMGDMGPDGDSGSGAGEGAA 181
 DB 64 ACRQPRKPIIEGDIYQDVDKSSRAVAFNL----- 98
 QY 182 PGAGAAAGSAGAFAGDGAAGGAAAPLSPGATVALLPAPPEFLWAFGLAKAGLRT 241
 DB 99 -----HSLKKGDIWALLMSNEPDPFVHWFGKLGCYV 132
 QY 242 AVFPLARRGPLHLRSCGAFALVAPPELSELPDLPALRMLHMAAGPTHPAGI 301
 DB 133 AFLNNIRNSNLLNCRACGPRALVAGADLLGTVEILPST-SENIYWGKMSV-PGV 190
 QY 302 SLLAEVSAEVDGPVGYLSSPOSTDCLYFTSGTGLPRAKISHKILQCGFYQL 361
 DB 191 ISLKERLSTSPDEPRSHHVSILKSTCLYFTSGTGLPRAKISHKILQCGFYQL 250
 QY 362 CGVHGEDVLYALPYHNSGLIGVCGNGIGATVYLKSKSAGOFMEDCOHRYTFQY 421
 DB 251 FCGTADDIYITLPLTHSSAAILIGSCVELATGATVLLKKKSASQFMSDCKYDVTYQY 310
 QY 422 ISELRYLYNOPPSAERGHKRYLAVSGLRPDMERFVRRGPOVLETGLEGVAT 481
 DB 311 ISELRYLYCKQKREBKHKRLAINGIRSDVREFLDPRGNKVELYLAITSSISF 370
 QY 482 INVTGKRGAVASWLYKHIFPFSIRYDVTIGBIRDPQCHMATSDEPGLVAPVSQ 541
 DB 371 KMYTGKIGALGTNLFYKLLSTFDLKYPQDEPMRNOGCIHVKKGEPGLISRYNA 430
 QY 542 QSPFLGAGPELAQCKLNDVFRGVDVFNTGDLJNDDOGLAFHRTGDTFRWKEN 601
 DB 431 KMPFGYAGPYHKTOKLCDVFKKGDVLTGDLIVDDNPLYFMDRTGDTFRWKEN 490
 QY 602 VATTAEAEVFEALDFOEVNVYGVTPGHEGAGMAALVLRPPHALDMQLYHVENLP 661
 DB 491 VATTAEADVIQMLDFIQEANYGVALISGEGRAGMAAIIKNTSLDEKYEQVTLPL 550
 QY 662 PVARPPFLQSLATTTFFKQKVRMANEGDPSTLSDPLVTDQAVGAILPLTARYS 721
 DB 551 AVACPFPLIOEKMENTGTFKLLKQLVEDGFNPLKISEPLVEMDKSYVLLRELVD 610
 QY 722 ALLAGLRI 730
 DB 611 QIMLGRIKL 619
 RESULT 10
 O9:WV6 PRELIMINARY; FRT; 620 AA.
 AC O9:WV6
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Solute carrier family 27 (Fatty acid transporter), member 2.
 GN SLC27A2.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Steinberg R.;
 RT "Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.
 RL EMBL; BC013442; AAH13442.1;
 DR EMBL; BC022170; AAH22170.1;
 DR EMBL; BC024735; AAH24735.1;
 DR MGI; MGI:1347099; SLC27A2.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00501; AMP-binding; 1.
 DR PROSITE; PS00455; AMP-BINDING; 1.
 DR SEQUENCE 620 AA; 70422 MW; 62994BDB1D828B37 CRC64;
 SQ
 Query Match 35.0%; Score 1346; DB 11; Length 620;
 Best Local Similarity 39.5%; Pred. No. 3.9e-84;
 Matches 275; Conservative 124; Mismatches 215; Indels 82; Gaps 7;
 QY 40 VPREEAGSMALLILPLIL--LLPLLLKLHMPQLRWLPADLAPVRLCKKALRA-- 95
 DB 2 LPVLYTG-LAGLILPLILITCCCPYLLQVRYFLRL--ANWARRVRSYRQRPVRLIL 56
 QY 96 RA-LAAAAADPEPBGCCSLARLAEALAQRAHFTLIGSRFSESEAESNRAPAF 154
 DB 57 RAFLQARRTKHP-----FLPRDETLYAQVDRSNQVARRL 95
 QY 155 LRALGMDWGDGDSGSGAGEGERAPAGDAAAGSAGAFAGDGAARAGGAAAPLSPG 214
 DB 96 HDQLG-----LRG 104
 QY 215 ATYALLPAPPEFLWAFGLARGLTAVPPLARRGPLHLRSCGAFALVAPPELES 274
 DB 105 DCVALFMGNEPAAVYTWLGLKLGCPMACLNINIRKSLHFCQCGAVLLASPDQRA 164
 QY 275 LEPDLPALRAMGLHMAAGGTHPAGISDLAEVSAEVDGPVGYLSSPOSTDCLYTF 334
 DB 165 VEEVPLTKDAVASVYVSTNTGNDVDTILKDVGSVSAEPPEBSRSSVETTPAVIYL 224
 QY 335 TSGTGLPRAKISHKILQCGFYQLCGVHGEDVLYALPYHNSGLIGVCGNGIGA 394
 DB 225 TSGTGLPRAKISHKILHRLWYGTGLAMSSGITAQDVITTMPLYHSAALMIGLHCIVGA 284
 QY 395 TVYLKSKESAGCFMEDCOOHRYTVFPQYIGELCRYLNOPPSAERGHKRYLAVSGLRPD 454
 DB 285 TLALRSKFSASQPWDCRKNVTVIYIGELRYLNTQKRPDRHKYKXLLGNGLRSD 344
 QY 455 TWRFVRRPGLQVLETYGLTEGNVATINYTGQKRGAVASWLYKHIFPFSIRYDVTG 514
 DB 345 VMREFIKRFDIHVEFYASTEGNIGFVNYPRKIGAVGRANYLQKVARLEYLIKVDKED 404
 QY 515 EPIRDPQCHMATSDEPGLVAPVSQSPFLGAGPELAQCKLNDVFRGVDVFNTGDL 574
 DB 405 EPRDANGYCIKPKKEVGLVCKITQLPFGYAGAGKQTKKRLDRFKKGDYFNSG 464
 QY 575 DLVCDQGLRPHDRTGDTFRWKENVATTEAEVFEALDFOEVNVYGVTPGHEGRA 634
 DB 465 DLMIDRENIVYHDXVGTDFRWKENVATTEADVLGVDFVEEVNVYGVTPGHEGRI 524
 QY 635 GMAALVLRPPHALDMQLYHVENLPYARPPFLQSLATTTFFKQKVRMANEGDP 694
 DB 525 GNASLKIKENYENGKLFQHLAEVLPASARPPFLNIQOTITGTFKIRKVTLMEEGSE 584
 QY 695 PSTLSDPLVTDQAVGAILPLTARYSALLAGLRI 730
 DB 585 PVIKDTLYFMDAKETPVPMTENTNAILIIDKIKL 620

RESULT 11

OBAYCS PRELIMINARY; PRT; 625 AA.
AC OBAYCS;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Similar to solute carrier family 27 (Fatty acid transporter), member 2 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_Taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strusberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC041746; AAH41746.1; -
DR GO: GO:0003824; F: catalytic activity; IEA.
DR GO: GO:0008152; P: metabolism; IEA.
DR InterPro: IPR00873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 2.
DR PROSITE: PS00455; AMP_BINDING; 1.
FT NON TER 1
SQ SEQUENCE 625 AA; 71253 MW; E5C22101B2DE51B4 CRC64;

Query Match 34.2%; Score 1312.5; DB 13; Length 625;
Best Local Similarity 39.5%; Pred. No. 7.8e-82;
Matches 268; Conservative 108; Mismatches 234; Indels 69; Gaps 6;
QY 54 LPPLLLPLLLKL-HLMPQLRMLPADLAFAVRAALCKCKALRAALAAAADEGEGG 111
DB 14 LPLGLLLILISFIPIFYFQJAWITAVRGIRI-RSSVSTPHATVVD----- 62
QY 112 CSLARLAEALAQQAHTFLIHGSRFSYSEMERSSNPAAPFLALGMDMGPDGSGE 171
DB 63 -----MFLKVERHDKPFVLEBEVYVYSHNDKLSNQAARA-DR----- 101
QY 172 GSAGGGEPAAPAGAPAAAGSGAEFAGDGAARGGAAAPLPGATVALLPAGPEFLTW 231
DB 102 -----KRAATKSGC-----VALFMANAPAYITW 126
QY 232 PGLAVALGRTAFVPTALRRGPLHCLSCGAPALVLAPELESLEPDLPALRANGLHMA 291
DB 127 LGVAALGSGIACLNINIRSGSEFLHCFRCRAKAVILAEPELKDVIIEVMPBLKXNVKVF 186
QY 292 AGPGHHPAGISDLAEVSAVDGPPVGLSSPQITDLCIYFTSGTGLPKAARISLTK 351
DB 187 LTKYVISEETESFLDKVRAADSEVPKGRFVSCKSLAMTYISGTTGLPRAALINHYR 246
QY 352 ILCCGFYQLCGVHEDVITYALPLTHMSGSLIGIVGCGIGATVLSKFSAGQFWDG 411
DB 247 ILSACGMEIICKVARADVYSPPLPYHSSAMMIGHGICSGATVLAKPKFSAGQFWDG 306
QY 412 QOHRVTVOYIGELCRVLYNOPPSKABRGHKRLAVSGSLRPDTERVVRPGLQVLET 471
DB 307 RRYNTVILYIGEVRLYLCNVPKSDDDVAHNVKMAIGGLRTDVWSEFLRRFGELHNEF 366
QY 472 YGLTGNVATINYTGQRGAVGASWLYKHIFPFLIRYDVTGEEIPDPQCHMATSPE 531
DB 367 VASTGNTAFITNYNTVSGVGRFDKILHSYDIKTDIEMDEVVRAMKCMKARAGQ 426
QY 532 PELLVAAPVSQSPPLGYAGPELAQGLKLDVFRGDVFNUTGDLVDDGGLRFDHRT 591
DB 427 PELLCKINSMPFGYAGDEHSTERKIMRDVFRKGDAYFNSGDLITVDQONFIYFHDV 486
QY 592 GPTFMKGENVATTEAAVEFALDPLQEVNVYGVVPGHEGAGAAALVLRPHALDMO 651
DB 487 GPTFMKGENVATTEAADICTIVNFIQEVNVYGVVQNHGRIGRAALILDEEVEFGDK 546
QY 652 LVTHVSENLPPYAPRFLQESLATTETFKQKVRMANEGDPSTLSDPLVLDQAVGA 711

DB 547 LVNAVHDFLPNVAAPRFFIRQNSMDITGTFKQKVLVGEFPAALISDPLVLDREKK 606
QY 712 YLPPLTARISALIAGNLRI 730
DB 607 YPMTQTIVEDIQKRIKL 625

RESULT 12

Q9Y2P5 PRELIMINARY; PRT; 690 AA.
ID Q9Y2P5
AC Q9Y2P5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Very long-chain acyl-CoA synthetase homolog 2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99410693; Pubmed=10479480;
RA Steinberg S.J., Wang S.J., McGuinness M.C., Watkins P.A.;
RT "Human liver-specific, very-long-chain acyl-coenzyme A synthetase: cDNA cloning and characterization of a second enzymatically active protein."
RT Mol. Genet. Metab. 68:32-42(1999).
RL EMBL, AF064255; AAD29444.1; -
DR Genew; HGNC:10399; SLC27A5.
DR GO: GO:0005783; C: endoplasmic reticulum; TAS.
DR GO: GO:0004467; C: long-chain-fatty-acid-CoA ligase activity; TAS.
DR GO: GO:0000038; P: very-long-chain fatty acid metabolism; TAS.
DR InterPro: IPR00873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP_BINDING; 1.
SQ SEQUENCE 690 AA; 75384 MW; 011313424D794546 CRC64;

Query Match 33.9%; Score 1303.5; DB 4; Length 690;
Best Local Similarity 42.1%; Pred. No. 3.7e-81;
Matches 297; Conservative 93; Mismatches 251; Indels 65; Gaps 13;
QY 28 GSGMFAAGM-NQVVP--IEAGSMALLLLPLLLPLLLKHLMPQLRMLPADLAFAV 84
DB 47 GLAMLARPMGVPVPPGLSLAAALALTLTP-----ARLPPGLRMLPADVITLA 95
QY 85 RALCKRALRAALAAAAADPEGEGCSLWRLAEALQQAARTFLIHGSRFSYSAE 144
DB 96 KILHLGUKR-----GC-----LSRQPDPTVDVAFERR-----AR 125
QY 145 RESNRAARALRLALGMDWGDGDDSGEGSANGSEERAPAGADAAAGSGAEFAGDGAAG 204
DB 126 AQGGERA-----LVVT-GP-----GAGSVTFEILDA-RACQAAWLKLEL-GDPAASIC 170
QY 205 GGAAPLSPGATVALLPAGPEFLTMFGIAXGLRTAFVPTALRRGPLHCLSCGARA 264
DB 171 AG-----EPTALLVLSQAVRA-LCMVLGIAGCTAMNPNRGRMLAHSVLSSGARV 224
QY 265 LVLAPEFLESLEPDLPALRANGLHMAAGGTHPAGISDLAEVSAVDGPPVPGYISSPQ 324
DB 225 LVVDPDLRESLEIILPKLQENIRCFYLSHTSPVGAALGALDAAPSHPPADIRAGI 284
QY 325 SIIDTCIYFTSGTGPKXARIISHLKIIQCGGFYQLCGVHEDVITYALPLTHMSGSL 384
DB 285 TWSPALFITSGTGLPKAALITHERVLDMSKMLSGATADVVYTYLPLVHWGLVY 344
QY 385 GIVGCGIGATVYLKSKFSAGQFWDGQOQRTVVOYIGELCRVLYNOPPSKABRGHKYR 444
DB 345 GILGCLDLAGTCLAPKFSSTSCFWDQRGVTVILYVGSILRYLCNIPQOPEDRTHYR 404
QY 445 LAVSGLRPDTMRFRRRPPLOVLETYGLTEGNVATINYTGQRGAVGASWLYKHIFP 504
DB 405 LAMNGLRADVWETFPQRFGRIRIWEYVGSYISNMGVLVNVGCGALGKQSCILRYLSP 464

Oy		505	SILRVDTTTEPRDRPOGCMATSPPEPGLVAPVSQSPFLGAGGELALOGCLLDVF	564
Dd		465	EIVGFPMDEANEPFRDNQFCIPVGCEPBELLTKTVSQOPFVGVRGPRESERLVAHNR	554
Oy		565	PPGDVPFNITDILLVCCDDQGFLEPHRHTGTPTFKKCNENATTEVAEVEFALDFLOEVNNG	624
Dd		525	QSQGVYNNNTDVTIAMREGEFLFYRRDLGTFPKMKNBSTHEBEGVSQVDFFLOOVNNG	584
Oy		625	VTVPGHGGRAGMALVLRPPHALDMQLTYHSENLPPYARPRFLQESLATEPFKKOQ	684
Dd		565	VCPDGCGKGKMAAVOLACQFTFDGSKLYQHRAALPAVAPHPIRIODAEVISTFKLM	644
Oy		685	KVRMANEGPDPESTLSPLLYLVDAVGAYLPILTARYSLLAGNLTIRI	730
Dd		645	KTRLVREGFNVGIYDPDLFVLIDNRASFRPLTAEMYQAVCCEGWRL	690
RESULT 13				
ID	088694		PRELIMINARY;	PRT; 662 AA.
AC	088694;			
DT	01-NOV-1998 (TREMBLrel. 08, Created)			
DT	01-AUG-1999 (TREMBLrel. 11, last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, last annotation update)			
DE	Fatty acid transport protein 5 precursor (PATP5) (long-chain fatty acid transport protein 5) (very-long-chain acyl-CoA synthetase related protein).			
GN	SLC27A5 OR VIACSR.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_Taxid=10099;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9837965; PubMed=96711728;			
RA	Hirsch D., Stahl A., Lodish H.F.;			
RT	"A family of fatty acid transporters conserved from mycobacterium to man.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:8625-8629(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BAB/C; TTSSUB=LIVER;			
RX	MEDLINE=98308102; PubMed=9642112;			
RA	Berger J., Tuppe C., Neumann H., Fares-Pelzer S.;			
RT	"A novel relative of the very-long-chain acyl-CoA synthetase and fatty acid transporter protein genes with a distinct expression pattern.";			
RL	Biochem. Biophys. Res. Commun. 247:1255-1260(1998).			
CC	-1- FUNCTION: INVOLVED IN TRANSLLOCATION OF LONG-CHAIN FATTY ACIDS ACROSS THE PLASMA MEMBRANE. MAY PLAY A PIVOTAL ROLE IN REGULATING AVAILABLE LONG-CHAIN FATTY ACID SUBSTRATES FROM EXOGENOUS SOURCES IN TISSUES UNDERGOING HIGH LEVELS OF BETA-OXIDATION OR TRIGLYCERIDE SYNTHESIS.			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PLASMA MEMBRANE.			
CC	-1- TISSUE SPECIFICITY: LIVER, BUT NOT IN FETAL LIVER. LOW LEVELS IN BRAIN, LUNG, TESTES, SPLEEN, AND SKELETAL MUSCLE.			
CC	-1- SIMILARITY: TO OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT COVALENT BINDING OF AMP TO THEIR SUBSTRATE.			
CC				
DR	EMBL; AF072760; AAC40189.1; -			
DR	EMBL; AJ223959; CAAL1688.1; ALT_INIT.			
DR	MGJ; MGJ:1347100; SLc27a5.			
DR	GO; GO:0016021; C:integral to membrane; IEA.			
DR	GO; GO:0003824; F:catalytic activity; IEA.			
DR	GO; GO:0006669; P:lipid transport; IEA.			
DR	GO; GO:0008152; P:metabolism; IEA.			
DR	InterPro; IPR000873; AMP-bind.			
DR	Pfam; PF00501; AMP-binding; 1.			
DR	PROSITE; PS00455; AMP BINDING; 1.			
KW	Glycoprotein, lipid transport; Transmembrane; signal; Transport.			
FT	SIGNAL	1	25	POTENTIAL.
FT	CHAIN	27	662	FATTY ACID TRANSPORT PROTEIN 5.
FT	TRANSMEM	28	48	POTENTIAL.
FT	TRANSMEM	156	176	POTENTIAL.

FT	TRANSMEM	312	332	POTENTIAL.
FT	CARBOHYD	533	533	N-LINKED (GLCNAC...) (POTENTIAL)
FT	CONFLICT	1	1	M -> MGWKKLTLLTLLTLLLVGLGQPPWPAAM (IN REF. 2)
FT	CONFLICT	61	61	K -> I (IN REF. 2)
FT	CONFLICT	381	381	T -> N (IN REF. 2)
FT	CONFLICT	541	541	C -> S (IN REF. 2)
FT	CONFLICT	661	661	N -> K (IN REF. 2)
SEQ	SEQUENCE	662 AA;	73251 MW;	1D7BAFE29FD2712 CSC64;
Query Match		32.6%	Score 1251;	DB 11; Length 662;
Best Local Similarity		39.8%	Pred. No. 1.4e-77;	
Matches	280;	Conservative	91;	Mismatches 255; Indels 78; Gaps 8;
QY	46	GSVAALLPLPLLL-----LPPLLLKLTLMPPOLRMPLADLFAVRA	86	
DB	10	GDPPLCVTLGLALLGRPMWISSMMHMLTSLVGAALTFLPLQPPGLRMHLKDVAFPMK	69	
QY	87	LCCRRALPAPALAAAADPEGEGCGSLAMFLAELAQQAANTLTLHGSRFFSYSEARE	146	
DB	70	LFYGLKRRRR-----LNKKRPP-----TIV-----DALRRQ	95	
QY	147	SNRAABAFLRALGMDWGPD---GDSGSGAGEGERAPAPAGCAAAAGAFAGCGDAA	202	
DB	96	-----ALAA---PDRVALVCTGSESSSTINSQDARSQDAWVTKKLDAVIQN	142	
QY	203	RGCGAAAPLSGATVALLLPAPREFLMLFGLAAG:RTAFVPTALRRGPLLCLRSQGA	262	
DB	143	TRDAALITVPSKTIASLS-----VFGLGATLQCPVAMINPHSGMPLSHSVRSQA	194	
QY	263	RAVLAPREFLESTEDPLPALRAMGLHMAAGPGNHPIGSDLLAEVSAEVDGAPVGYLSS	322	
DB	195	SVLLVDPDLQNLLELVLPKLAENIHCGYLAHSSPTTGVEALGASLDABSDVPASLIRA	254	
QY	323	PGSITIDICLVYFTSGTGTLPKPARISHKILQCCGFYQLCGVHDEVIYIALPLPYHMSG	382	
DB	255	TIKKSPALIFFTSGTGLPKPALLSHBRVLYQSVNLSFGCGRADVVYVDVPLPYHTIGL	314	
QY	383	LLGIVCGMGIGATVYLKSKFSAGGFMEDCCQNRITVYQYIGELGRYLMNQPSKAERGHK	442	
DB	315	VLGLFGCIQVQATVLPKPSASAFMBACRGGHGVTLVYGEILIRYLCNWEDEDEDKHT	374	
QY	443	VLAVVSGLRDPTWERFVRFRGPQVLEVYGLTEGNVATINITYGQRCAGVGRASMLYKHI	502	
DB	375	VLAMGTLGRANWKNQOQRGRFIRIMEFYGSTEGNGLMNVYHGCAVGRISCLIRMLT	434	
QY	503	PFSILIRVDVTGERTIRDPQCHMA:TSQSPRLVLA:VYSQSPFLGYAGSPRLAOGKLLKD	562	
DB	435	PFEVLQPDIEFAEFLRKQGCFCIVEGSKQELLTKTKRKQPFLLGYGSGQAESNRKLVAN	494	
QY	563	VFRPGDVAFPMNGDLLVDDDDGFLFPHRTGTGTFPMKENVATTEVAEFEALDPLQEVNV	622	
DB	495	VRRVGDLYFNTGDTLTDQGBEFFFDQBLGDTFFMKGENNSTGEVECVLSLDDLEEVNV	554	
QY	623	YGVTVPGHGRAGNAALVLRPPHALDLMQVLYTHVSENLPYARDFRRLQESLATTETFK	682	
DB	555	YGVVPVPGEGKVGMAAVKLAPAGTKTFDDQKLYQHFRSMLPAVATPHFIRIQDSLEITWYK	614	
QY	683	QQRKMANEGRPSTLSDPLVLDQAVGAYILPTTAYYSALLAG	726	
DB	615	LKRSRLVREBGFVGIIADPLVILDNKAQTFERSLMPDVYQAVCEG	658	
RESULT 14				
091VDS				
ID	091VDS	PRELIMINARY;	PRT;	689 AA.
AC	091VDS;			
DT	01-DEC-2001	(TREMBLrel. 19, Created)		
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25, last annotation update)		
DE	Hypoetical protein (Solute carrier family 27 (Fatty acid transporter), member 5).			
GN	SLC27A5.			

```
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC013335; AAH13335.1; -.
DR EMBL; BC013272; AAH13272.1; -.
DR PIR; J0107; J0107.
DR MGI; MGI:1347100; S1C27A5.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KM Hypothetical protein.
SQ SEQUENCE 689 AA; 76202 MW; 1642BBC2CF04FAA3 CRC64;

Query Match 32.6%; Score 1251; DB 11; Length 689;
Best Local Similarity 39.8%; Pred. No. 1.5e-77;
Matches 280; Conservative 91; Mismatches 255; Indels 78; Gaps 8;

QY 46 GSMALLLLPLLL-----LPLLLKLHLPOLRLPADLAFVRA 86
DB 37 GDPPLVLVLGLALGRPWISSMFWMLSLVGAALTLFLPLQPPGLRLMKDVAFTTKM 96
QY 87 LCCRKALPAPALAAAADPEGEGGCSLAWRLAELAQRAAHTFLHGSRRFSYSEARE 146
DB 97 LFYGLKFFRR-----LNKHPPE-----TV-----DALERQ 122
QY 147 SNRAARAFRLALGWMDGPDG---GDSGSGSAGEGERAPAGDAAGSAGAFAGCGGA 202
DB 123 -----ALAW---PDRVALVCTGSESSSTNSQDARSQDAWYLRKKLMDAVIQN 169
QY 203 RGGGAAPLSPGATVALLPAGPEFLWMLFGLAKAGLTAFTALRRGPLHLCRCGA 262
DB 170 TRDAALIVLPSTKISALS-----VFLGAKLGCVPAMINPHSRGMPLHSVRSQA 221
QY 263 RALVLAPEFLSEBDLPALRAMGLHMAAGPCHPAGISDLAEVSAEVNPGPGLSS 322
DB 222 SVLIYDPLQENLEEVLPKLAENHCHFYLGHSSTPGVEALGSLDAAPSDPVASLRA 281
QY 323 POSITDCLVYFTSGTGLPKAARISHLKIQCQGYQLCGVHOEDVYIALPYHMSGS 382
DB 282 TIKWKSPIAFITSGTGLPKALISHERVIOVSNVLSFCGRADVDVYDLPYHTIGL 341
QY 383 LLGIYGCWIGATVVLKSKFSAGQWEDCCQOHRVTVFOYIGELCRYLVNPPSAAERGH 442
DB 342 VLGFGICQVATVCLAPKFSASRFMAECROHGVVILYGEILRYLCNVPEOPEDKHT 401
QY 443 VRLAVSGLRPDVWERFVRFGPLOVLETYGTGEGNVATINYTGORGAVGASWLYHIF 502
DB 402 VRLANGNGLRANWKNFOQRFGPIRIWEPYGTSGENGLMNYVGHGCAVGTSCILMLT 461
QY 503 PPSLIHYDVTGPIRDPQGHCMATSPGEGILVAVPSQSGPFLGYAGPELAAQGLKLD 562
DB 462 PPELVQDFIETAEPLRDKQGFCEIPEVPGKPGILLTKYKKNQPLGYSGSQAESNRKLVAN 521
QY 563 VERPDVAFNMGDILVCDQDGFRLHDSGTDFRMKGENVATTEVAVFEALDLQGVNY 622
DB 581 VRRGVDLFFNIGDVLTLQDGGFFYQDLGDTFRMKGENVSTGEVCEVLSLDLLEEVN 581
QY 623 YGVTVPGHEGRAGMAALVLRPPHALDMQLYTHVSENLPYARPRFLRLQSLATTTTFK 682
DB 582 YGVVPVPGCEGKVMAAVLAPOKGTQDCKLVHVRSMPLPAAVYAPHFIRIQDLSLITNTYK 641
QY 683 QOKRMANEGPDPSTLSPLVYLDQAVAYPLTTARISALLAG 726
DB 642 LVKSLVREGEFDVGIADPLVILDNKAQTFRSLMPDYYQAVCEG 685
```

```
RESULT 15
Q9ES38 PRELIMINARY; PRT; 690 AA.
ID Q9ES38
AC Q9ES38
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Bile acid CoA ligase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Falany C.N., Xie X., Wheeler J., Wang J., Barnes S.;
RT "Molecular cloning and expression of rat liver bile acid CoA ligase.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF242189; AAG09770.1; -.
DR GO; GO:0016874; F: ligase activity; IEA.
DR GO; GO:0008152; P: metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP_BINDING; 1.
KM Ligase.
SQ SEQUENCE 690 AA; 76265 MW; D5BEB8317758DA59 CRC64;

Query Match 32.0%; Score 1231; DB 11; Length 690;
Best Local Similarity 39.2%; Pred. No. 3.6e-76;
Matches 280; Conservative 98; Mismatches 267; Indels 70; Gaps 12;

QY 52 LLPLLLPLLLKLHMP-----QLRWLPAD-----LAFVRLQCK-----RAL 93
DB 10 LLSLILLVG---LQPLPAAATALLRNFLGDPCTFVLGLAFLGRPWISSITPWLSL 66
QY 94 RADAALAAAADPEGEGGCSLAWRLAELAQRAAHTFLHGSRRFSYSEARESNR-AAR 152
DB 67 AAALITSLIPRPDP---ELRWLHSDVA---FAFLILEYG-----LNLRRRLNHPPS 114
QY 153 AFLRALGMDKPGDGGSGSAGEGERAPAGDAAGSAGAFAGCGGA 207
DB 115 LFLVDAL-----EQQAAPDQVALVCTGSESSSTNRELNARACQ 155
QY 208 AAPL-----SPGATVALLPFA-GPEFLWMLFGLAKAGLTAFTALRRGPLH 255
DB 156 AMALKKALKRATIQEDKATAILVPSKSIASLAVLGLAKLGCVPAMINPHSRGMPLH 215
QY 256 CLRSQARALVLAPEFLSEBDLPALRAMGLHMAAGPCHPAGISDLAEVSAEVDGP 315
DB 216 SVQSSASVLIYDPLQENLEEVLPKLAENHCHFYLGHSSTPGVEALGSLDAAPSDP 275
QY 316 VPGYLSPOSITPTCYIFTSGTGLPKAARISHLKIQCQGYQLCGVHOEDVYIALP 375
DB 276 VPKLKAANKKSPAFITSGTGLPKALISHERVIOVSNVLSFCGRADVDVYDLP 335
QY 376 LYHMSGSLGIYGCWIGATVVLKSKFSAGQWEDCCQOHRVTVFOYIGELCRYLVNPPS 435
DB 336 LYHSMGLVIGVGLCTCLAPKFSASRFMAECROHGVVILYGEILRYLCNVPEO 395
QY 436 KAERGHKVLAVSGLRPDVWERFVRFGPLOVLETYGTGEGNVATINYTGORGAVGRAS 495
DB 396 PEDKKTIVFPAIANGNGLRANWKNFOQRFGPIRIWEPYGTSGENGLMNYVGHGCAVGT 455
QY 496 WLYKHLPPSLIRYDVTGPIRDPQGHCMATSPGEGILVAVPSQSGPFLGYAGPELA 555
DB 456 CFIRMLTPLELVQDFIETAEPLRDKQGFCEIPEVPGKPGILLTKIRKNQPFLLGYRSGQDET 515
QY 556 QGKLLDVPFRPGVFNNTGDLVCDQDGFRLHDSGTDFRMKGENVATTEVAVFEALD 615
DB 516 KERLVANVQVQDGLVYNTGDLVLDQDGGFFYQDLGDTFRMKGENVSTGEVCEVLSLD 575
QY 616 FLQEVNVVGVTVPGHEGRAGMAALVLRPPHALDMQLYTHVSENLPYARPRFLRLQSL 675
```

Db 576 FLEEVNXYGVTVPEGCEGKVMGMAAVKLAPEGKTFDGGKLYOHVRSWLPAYATPHFIRIODSL 635
Qy 676 ATTEFFKQOQVRMANEGFDPSTLSDDLVLVDQAVGAVLPLTTARYSALLAGNLR 730
Db 636 EITNTYKLVKSQLAREGFDVGVIADPLYILNKKAFETFRSLMPDVYQAVCEGTWKL 690

Search completed: June 30, 2004, 17:34:58
Job time : 54 Secs

Best Local Similarity 39.7%; Pred. No. 8.6e-83;
Matches 276; Conservative 123; Mismatches 215; Indels 82; Gaps 7;

```

QY 40 VPIEAGSMALLLPLLL--LLPLLLKLLHMPQLRPLADLAVALRALKCKKALARA-- 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 LPVLVTG--LAGLLPLLLTCCCPYLLODVRFPLD-----ANMARFASVYRQRFPVPTIL 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 96 RA-LAAAADPEPGGCSLAWRLAEIAQAQAATFLIHGSRRSYSAEERESNRARAF 154
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 57 RAFLQAKRTPHK-----FLFRBETLTYAQVDRSRNQVVAL 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 155 LRALGMDWPGDGDGSGSAGEBAPAGDAAGAAAGSAPAGDGAARGGGAAPLSPG 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 HDQL-----LRQG 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 215 ATVALLPAGEPEFLMTFGLAKAGLRTAFVPTALRGPHLLHRCSCGARALVLAPELES 274
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 DCVALFPGNEPARYVIMIGLKLGGCPMACLNINIRAKSLHGFQCCGAKVLAPDDQEA 164
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 275 LEPDLPALRAMEGLHMAAGPGTHPAGISDLLAEVSAEVDGVPVGYLSPQITDTCLYIF 334
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 165 VEALPTLKDAVSVFVYSRTSNTNGVDIILDKVDGSAEPTPSWSEVFTTPAVYIY 224
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 335 TSGTGLRKAARISHLKLQCGFYQLGCVHEDVITYALPLMHSSGLIGVGMGIGA 394
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 225 TSGTGLRPAATINHRRLMYGTGLAMSSGITADVITYTTPPLYHSAALMIGLHGLIYGA 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 395 TVVLKSKFSAGQFWECDQOHRVTVFOYIGELCRYLVNPPSKAERGHKVLAVSGLRPD 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 TLAKSKTSASQFMDCKRYNTVTOYIGELRLYLCTPQKPNDRDEKVKALANGRGD 344
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 455 TIERVRRRGPLOVLETGTLTGNVATINYGQRANVGRASMLYKHIFPFLIAYDWTG 514
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 VIREPIKRGDILHYVEFYASTEGNIGFVNPRIKIGAVRANVLOKVARVYLIRYDEKD 404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 515 EPIRPOGCMATSGEPGLIAPVSOQSPFLYAGPELDAQKLLKDVFPFGDVFNTG 574
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 405 EPRVANGCICIVPGEHGLVCKITQLPFIYAGGKTQTEKKLRAVFKGDIYFNSG 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 575 DILVCDQGFLEPHRTGDTFRKGENVATTEVAEVALDPLQEVNYYGVTVGHGERRA 634
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 465 DILMTIDRENFVYFHRVGDTPFRKGENVATTEVADIVGLVFEVENVYGVVPGHEGRI 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 635 GAAALVLRPPHALDLMOLYTHSENLPYARPRPRLQESLATTETFRKQKVRMANEGFD 694
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 525 GNASLKIKENYFNGKLLFOHIAEYLPYARPRPRLQDITIEITGTFHKKVTLMEBGFN 584
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 695 PSTLSDPLVLDQAVGAYLPLTTARYSALLAGNLRI 730
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 585 PTYIKDTLYFMDAEKTFVPMTEINYNALIIDKTLK 620
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RT synthetase."/
RL J. Biol. Chem. 271:30360-30365(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
   family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, D85100; BAA12722.1; -
CC InterPro; IPR000873; AMP-bind.
CC Pfam; PF00501; AMP-binding; 1.
CC PROSITE; PS00455; AMP BINDING; 1.
CC Ligase; Fatty acid metabolism; Peroxisome; Transmembrane.
FT TRANSMEM 1 21
FT TRANSMEM 107 127
FT TRANSMEM 262 282
FT SEQUENCE 620 AA; 70693 MW; 6CF9362DC3805526 CRC64;
SQ

```

Query Match 34.9%; Score 1340; DB 1; Length 620;
Best Local Similarity 39.1%; Pred. No. 4.7e-82;
Matches 271; Conservative 123; Mismatches 223; Indels 76; Gaps 5;

```

QY 40 VPIEAGSMALLLPLLL--LLPLLLKLLHMPQLRPLADLAVALRALKCKKALARA 97
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 LPVLVTG--LAGLLPLLLTCCCPYLLODVRFPLD-----ANMARQVSRQRFRVR-- 53
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 96 LAAAADPEPGGCSLAWRLAEIAQAQAATFLIHGSRRSYSAEERESNRARAFARA 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 54 -----TILVPIEQAKRTPHKFLPRBETLTYAQVDRSRNQVAAALHDH 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 158 LGMDWPGDGDGSGSAGEBAPAGDAAGAAAGSAPAGDGAARGGGAAPLSPGATV 217
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 99 LG-----LRQGDV 107
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 218 ALPLPAGPEFLMTFGLAKAGLRTAFVPTALRGPHLLHRCSCGARALVLAPELESLEP 277
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 ALFPGNEPARYVIMIGLKLGGCPMACLNINIRAKSLHGFQCCGAKVLASPELHAYBE 167
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 278 DLPALRAMEGLHMAAGPGTHPAGISDLLAEVSAEVDGVPVGYLSPQITDTCLYITSG 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 168 VLPFLKEGVSVYVYSRTSNTNGVDIILDKVDGVSADPIFESWRSVFTTPAVYIYTSG 227
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 338 TTGLPRAARI SHLKLQCGFYQLGCVHEDVITYALPLYHNSGSLIGVGMGIGATV 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 TTGLPRAATINHRRLMYGTSLARSGIKAHADVITYTTPPLYHSAALMIGLHGLIYGAIFA 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 398 LKSKFSAGQFWECDQOHRVTVFOYIGELCRYLVNPPSKAERGHKVLAVSGLRPDTEWE 457
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 288 LKSKFSAGQFWECDQOHRVTVFOYIGELRLYLCTPQKPNDRDEKVKALANGRGDVMR 347
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 458 RYVRRRGPLOVLETGTLTGNVATINYGQRANVGRASMLYKHIFPFLIRDYTTGRI 517
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 348 EPIKRGDILHYVEFYASTEGNIGFVNPRIKIGAVRANVLOKVARVYLIRYDEKDEVP 407
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 518 RDQOCHMATSGEPGLIAPVSOQSPFLYAGPELDAQKLLKDVFPFGDVFNTGDL 577
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 408 RDANGCICIVPGEHGLVCKITQLPFIYAGGKTQTEKKLRAVFKGDIYFNSGDL 467
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 578 VCDQGFLEPHRTGDTFRKGENVATTEVAEVALDPLQEVNYYGVTVPGHEGERRA 637
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 468 MIDRENFIYFHRVGDTPFRKGENVATTEVADIVGLVFEVENVYGVVPGHEGRIIGA 527
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 638 ALVLRPPHALDLMOLYTHSENLPYARPRPRLQESLATTETFRKQKVRMANEGDPST 697
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 528 SIKMKENYFNGKLLFOHIAEYLPYARPRPRLQDITIEITGTFHKKVTLMEBGFNPSV 587
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 698 LSDPLVLDQAVGAYLPLTTARYSALLAGNLRI 730
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 2
ID VICS_RAT STANDARD; PRT; 620 AA.
AC P97584;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Very-long-chain acyl-CoA synthetase (EC 6.2.1.-) (Very-long-chain-fatty-acid-CoA ligase).
GN SLC27A2 OR FACL1 OR VLACS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Mistar; TISSUE=Liver;
RX MEDLINE=97094763; PubMed=8939997;
RA Uchiyama A., Aoyama T., Kamijo K., Uchida Y., Kondo N., Orit T., Hashimoto T.;
RT "Molecular cloning of cDNA encoding rat very long-chain acyl-CoA


```

Db      250  SGTGLPAAIVSVSRYYRIAIFGHHSYSKMANVLDCLPLYSANINMGVCIIYGL 309
Qy      395  TVVLKSKESAGQFWECCQHRVTAFQYIGELCRYLVNQPSPKARHGKVLAVSSGLRPD 454
Db      310  TVLKRKFSASRPFMDCKVKNCTVQYIGELCRYLVNQPSPKARHGKVLAVSSGLRPD 369
Qy      455  TWSEFVRARFGLQVLEYIGLTVGNVATNTYQGRANGRAMLYKHIFPSLIIYDVTG 514
Db      370  TWSEFTQGFVGRQIGEFGATENCSCILAMDKGSCGFSRILTHYFIRLVXNEDTM 429
Qy      515  EPIDPOGHCMATSPSEGLVLVAPQOSP---FLGYAGBELAQGLKLDVFPQGVVF 571
Db      430  EPLDSSQGLCTPCQGPBGLVQINQDPLRFDDGY-SISANKKIASVFKGSAY 488
Qy      572  NTGDLVLDQGFLEFHDRTGDFRMKGENVATTEVAEVEBALDPLGEVNVYGVTVGHE 631
Db      489  LSGDLVMDDELGYMFRDRSGDTPRMGENSTTEVAEVALSRLIGQDVAVGYAVGVGE 548
Qy      632  GRAGMAALVLRPPHALDMQLYTHVSENLPPYARPRFLRLOESLATTETPKQGVRRANE 691
Db      549  GKSGMDA-IADPHQLDPSNMYQELQKVLASACPILFRLPLQVDTTGTFRKIQKRLQRE 607
Qy      692  GFDPSTLSDPLVYLDQAVGAYLPLTARYSALLAGNLR 730
Db      608  GFDPQTSDRFLPFLDKQGRPLPDERVHARICAGDSSL 646

```

RESULT 6 FAT1_YEAST STANDARD; PRT; 623 AA.

```

AC      P38225;
DT      01-OCT-1994 (Rel. 30, Last sequence update)
DT      01-OCT-1994 (Rel. 40, Last annotation update)
DE      Probable long-chain fatty acid transport protein.
GN      FAT1 OR YBR041W OR YBR0411.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CX      Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=S288C;
RA      Andre B., Czechluch C., Hein C., Jauniaux J.C., Urrestarazu A.,
RS      Viessers S.;
RL      Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RD      CHARACTERIZATION.
RX      MEDLINE=97236810; PubMed=9079682;
RA      Paergerman N.J., Diruseo C.C., Elberger A., Knudsen J., Black P.N.;
RT      "Disruption of the Saccharomyces cerevisiae homologue to the murine
RT      fatty acid transport protein impairs uptake and growth on long-chain
RT      fatty acids."; 272:8531-8538 (1997).
RL      J. Biol. Chem. 272:8531-8538 (1997).
CC      -1- FUNCTION: May be involved in long-chain fatty acids uptake, and
CC      thus may play a pivotal role in regulating their accessibility
CC      prior to metabolic utilization. May play an important role in
CC      uptake of these hydrophobic compounds under conditions where fatty
CC      acid synthesis is compromised.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC      -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC      family.

```

```

CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC      or send an email to license@ebi.ac.uk).

```

```

DR      EMBL; Z35910; CAA84983.1; -.
DR      PIR; S45899; S45899.
DR      GenBank; U00000; -.
DR      SGD; S0000245; FAT1.
DR      GO; GO:0005811; C:lipid particle; IDA.
DR      GO; GO:0005792; C:cytosol; IDA.
DR      GO; GO:0005777; C:peroxisome; IDA.
DR      GO; GO:0005886; C:plasma membrane; IMP.
DR      GO; GO:0005324; F:long-chain fatty acid transporter activity; IMP.
DR      GO; GO:0004467; F:long-chain-fatty-acid-CoA ligase activity; IMP.
DR      GO; GO:0006869; P:lipid transport; IMP.
DR      GO; GO:0000038; P:very-long-chain fatty acid metabolism; IMP.
DR      InterPro; IPR000873; AMP-bind.
DR      Pfam; PF00501; AMP-binding; 1.
DR      ProSITE; PS00455; AMP_BINDING; 1.
KW      Lipid transport; Transmembrane.
FT      TRANSMEM 6
FT      TRANSMEM 54
FT      TRANSMEM 149
FT      TRANSMEM 293
FT      CARBOHYD 184
FT      CARBOHYD 289
FT      CARBOHYD 534
FT      CARBOHYD 591
SO      SEQUENCE 623 AA; 71697 MW; 0AE0270B60C8CFE CRC64;

```

Query Match 17.0%; Score 654; DB 1; Length 623;
Best Local Similarity 33.7%; Pred. No. 2.9e-36;
Matches 169; Conservative 84; Mismatches 193; Indels 56; Gaps 14;

```

Qy      214  GATVALLDPAGEPEFLIMFGLAKAGLRTAFVPTALRGLHCLRCGARALVLP--- 269
Db      137  GDVAIDCTNKLPLFVFLWLMSTNIGALPFAFLVNNKGTPLVSLKISNTQYFIPDASN 196
Qy      270  ---EFLESLEPDLPALRAMGLHMAAGPETHAPGASDLIAVSABVDGVPGLS---- 321
Db      197  PIRESESEIKNALPDVKNLVLE-----EQDLNHEH---LNSQSPFLQDNVR 241
Qy      322  SPOSIND--TCYIFPSGTGLPKAARISHLK-LIQQCGFYQLCGHQBQVYLLPLVH 378
Db      242  TPLGLDFRPSMLIYSGTGLPKSAIKSRKSSVGCQVFGVHLMTNISTVFTAMPLFH 301
Qy      379  MGSGLIGVCMKIGATVVLKSKFSAGQFWECCQHRVTAFQYIGELCRYLVNQPSPKAE 438
Db      302  STALLGACALISHGCLALSHKFSASTFWKQVYLGAHIGYVBSVCYLLHTPTISKYE 361
Qy      439  RGHKVTLAGSGLRPDPTWRFRFRGGLQVLEYIGLTVGNVATNTYQGRANGRAMLYKH 495
Db      362  KQHKVAVANGSRPDIWDFRFRFRIEVIYGFVTAATEAPFAITTF--QKGFPGIACRN 419
Qy      496  -----WLYKHIFPSLIRYDVTGEPF-RDPQGHCMATSPSEGLVLA---PVSQOSP 544
Db      420  YGTLIQLWFLS--FQGLVMDPNDSDVIRNSKGCCEVAVPGBPBLMRIRFPKPKETS 477
Qy      545  FLGYAGPELAAQGLIKLVFRPGDYFENTGDLVLDQGFLEFHDRTGDFRMKGENVAT 604
Db      478  FQGLYNNAETKSKVVRDVFRRGDAMWYRCGDLILKDEYLLWYFLDRMGDTFFMKSENVST 537
Qy      605  TEVAEVEBALDPLQ---EVNVYGVTVGHEGRAGMALVLRPPHALDMQLYTHVSE--- 658
Db      538  TEVEDQLTNSKEQYAVQVAVIKVFKYGRGRGFAVIKL-TNSIDITAKTLLNDSLSR 596
Qy      659  NIPYARPRFLRLOESLATTTE 679
Db      597  LNDPSYAMPPLFYFVDEIKMTD 618

```

```

RESULT 7
CAIC_EC057 STANDARD; PRT; 522 AA.
ID CAIC_EC057
AC Q8XA34;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

```

DT 28-FEB-2003 (Rel. 41, Last annotation update) (EC 6.3.2.-).
 DE Probable crotonobetaine/carnitine-CoA ligase (EC 6.3.2.-).
 GN CAIC OR 20043 OR EC50040.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 RN NCBI_TaxID=83334;
 RP SEQUENCE FROM N.A.
 RC STRAIN=0157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel E.J., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimallanta E.T., Potamoustis K.,
 RA Apodaca J., Mantharman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.,
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7",
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=0157:H7 / RIMD 0509952;
 RX MEDLINE=21156331; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasekawa C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.",
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: Could catalyze the transfer of CoA to crotonobetaine
 CC or carnitine (By similarity).
 CC -1- PATHWAY: Carnitine metabolism (conversion of carnitine to
 CC gamma-butyrobetaine).
 CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as there are no restrictions on its
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE005180; AAG34340.1; -;
 DR EMBL; AP002550; BAB33463.1; -;
 DR PIR; H85484; H85484.
 DR PIR; H90633; H90633.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KW Ligase; Complete proteome.
 SQ SEQUENCE 522 AA; 59144 MW; D0F0AB1488B7D8C9 CRC64;
 Query Match 10.6%; Score 405.5; DB 1; Length 522;
 Best Local Similarity 25.8%; Pred. No. 8.9e-20;
 Matches 154; Conservative 78; Mismatches 235; Indels 131; Gaps 19;
 QY 110 GGCSLAWRLAELAQQAATPLIHGS-----RRFSYSAEESNRAPAFRLALGMDWGP 164
 DB 10 GGCHLRQWMDLADYGHKHALICSSGGVNRYSYELNDEINKTANLFF-TLG----- 63
 QY 165 DGGSDGSGSAGEGAAAGAGDAAAGSAGAFAGCGARGGGAAPSPGATVLLPAG 224
 DB 64 -----IKGDKVALHLDNC 77
 QY 225 PEFLLWTFGLAVAGRTAFVPTALRGLLCLRCGARAFLVLA-----PEFLSIEPLLP 280
 DB 78 PEFICWFGLAKIGAIMVPINARLLRBSAWLTIQNSQCLIVTSQGFPMYQIQCOEDAT 137
 QY 281 ALRANGHLHMAAGPETHRAGISDLLAETSAEVDGVPFY-----LSSPGSTTD 328

DB 138 QLR-----HI-----CLTDVALPADGVSSEFTOLKNOQAPATLCYAPLSTDD 179
 QY 329 TCIYITSTGTGLPKRARIISHUKIIQCGFYOL--CGVHOEDVYLALPLYMSSGLGI 386
 DB 180 TKEILFTSPTRPKGVITHTN-LRPAGYSAMOCALDDVVITLWPAFIIDQCCTRA 238
 QY 387 VCGMGIGATVVLKSKESAGQFWECCQHRVYFOYIGELCRYLVNPPSKARGRKVR-L 445
 DB 239 MAFSAGATFVLVEKYSARAFWQGVQKXARATITECIPMMIRLTMTWQPSANDRRLREV 298
 QY 446 AUGSGIRPDTWRFPARRFPLOVLETYGITEGNVATINTGQ-----GAVGSASMLY 498
 DB 299 MEYLNLSBOEKPTFCERFG-VLLTSYGMTETVGI--GRRPDKRWPISIGAGFY 354
 QY 499 KATFPSSLIRYDVTGEPFRDPQGHQWATSPGEPG-LVAPVSCQSPFLVAGGDELAQG 557
 DB 355 -----DAEIRD--DINRPLPAEIEICXGVPKGTIFREYILNPK-ATA 396
 QY 558 KLIKDFRPQDVFEHTGDLIVCDQGLRFLHRTGDTFRMKEENATTEVAEVPALDPL 617
 DB 397 KYLE-----ADGMLHGTGTYRDERGFFYIDRCNMTKGGENVSCVELENIATHPKI 451
 QY 618 QENVNYGVTVPGHEGRAGMAALVLRPHALDMQYTHVSENLPYAPRFRLOEST 675
 DB 452 QDIIVVGIR-DSIRDEALRAFPVLNGETLSEEEFFRCEQMAKFPYSYLEIKDL 508
 RESULT 8
 CAIC_ECOLI STANDARD; PRT; 522 AA.
 ID P31552;
 AC 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable crotonobetaine/carnitine-CoA ligase (EC 6.3.2.-).
 GN CAIC OR B0037.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RP STRAIN=04:K74;
 RX MEDLINE=95115548; PubMed=7815937;
 RA Eichler K., Bourgis F., Buchet A., Kleber H.-P.,
 RA Mandrand-Berthelot M.-A.;
 RT "Molecular characterization of the cat operon necessary for carnitine
 RT metabolism in Escherichia coli.",
 RL Mol. Microbiol. 13:775-786(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizobuchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 0-2.4 min region.",
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of Escherichia coli K-12",
 RL Science 277:1233-1238(1997).
 CC -1- FUNCTION: Could catalyze the transfer of CoA to crotonobetaine
 CC or carnitine.
 CC -1- PATHWAY: Carnitine metabolism (conversion of carnitine to
 CC gamma-butyrobetaine).
 CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme

CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X73904; CAAS2113.1; -
 DR EMBL: D10483; BAB96606.1; -
 DR EMBL: AE000114; AAC73148.1; -
 DR PIR: B64724; S40558.
 DR HSSP: P08659; 1LC1.
 DR Ecogen: EG11558; calC.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KM Ligase; Complete proteome.
 FT VARIANT 103 103 C -> R (IN STRAIN O44:K74).
 FT VARIANT 257 258 RA -> PR (IN STRAIN O44:K74).
 FT VARIANT 291 291 Q -> R (IN STRAIN O44:K74).
 FT VARIANT 333 333 I -> S (IN STRAIN O44:K74).
 FT VARIANT 350 350 V -> A (IN STRAIN O44:K74).
 FT VARIANT 379 379 I -> V (IN STRAIN O44:K74).
 FT VARIANT 393 393 Q -> K (IN STRAIN O44:K74).
 FT VARIANT 401 401 A -> V (IN STRAIN O44:K74).
 FT VARIANT 413 413 R -> C (IN STRAIN O44:K74).
 FT VARIANT 417 417 D -> G (IN STRAIN O44:K74).
 FT VARIANT 447 447 A -> T (IN STRAIN O44:K74).
 FT VARIANT 508 508 L -> LDI (IN STRAIN O44:K74).
 SQ SEQUENCE 522 AA; 59089 MW; 1C54D003B130671 CRC64;
 Query Match 10.2%; Score 393.5; DB 1; Length 522;
 Best Local Similarity 25.7%; Pred. No. 5.6e-19;
 Matches 155; Conservative 79; Mismatches 227; Indels 143; Gaps 22;
 QY 110 GGCSIALMLALAOQRAHTPLHGS-----REFSYSAESESRRARAPLALGWMGP 164
 DB 10 GGQHLRQWMDLADVYGHKTLALICSSGGVNVRSYLELNEIRTNLFY-TLG----- 63
 QY 165 DGDGSGESAGEGERAAPGADAAAGSAEFAAGDGAAGGAAPSPGATVALLPAG 224
 DB 64 -----IRKGDVXVHLHNDNC 77
 QY 225 PEPLMLWGLAKAGIRTAFTPTALRGGPLHC-----LNSCGARAVLA---PEFLS 274
 DB 78 PEPLFCWGLAKIG--ALMVPINR---LLCESAWILQNSQACLVTSAQFYPMYQOI 131
 QY 275 LEPLPALRAMGLHMAAGPGTHPAGISDLAEVSAEVDGVPVQY-----LSS 322
 DB 132 QOEBAIQIR---HI-----CLTDVALPADDDVSSFTLQKXQOATLCTYAP 173
 QY 323 POSTTDTCLVFTSGTGLPRAARISHUKILQCGFYQL--CGVHOEDVIYALPLYMS 380
 DB 174 PLSTDDTDLEILFTSGTTRPKGVVITHYN-LRPAQYSAWOCALRDDVDVLTVPAPHD 232
 QY 381 GSLIGIVCGMIGTAVYLKSKFSAGQFEDCCQHRVYFYQYIGLCLVYLVNQPSKXERG 440
 DB 233 CQCTRAWMAFASGATFVLEKYSARAFWQYQKTRAVTETCPMIRLWQVPSANDQO 292
 QY 441 HKVR-LAVSGLRPDTWERFVRRGPGQLVETVGLTEGNVATINYTQGR-----GAVG 492
 DB 293 HRLRREVMYVLYLSEGEKACFCERFG-VALLTSYGTETIVGII---GDRPEDKRWPSIG 348
 QY 493 RASWYIKRIFFPSLRIVYVTGEPFRDQGHCMATSPERP-LVAPVSOQSPLGYAGG 551
 DB 349 RVGFEYE-----AEIRD--DHNRLPAEIEIGIKIGPGKTIKFEFLN 391
 QY 552 PELAGKLLKDFRGDFVFNVTGDLVCDQGFTRFHRTGDTFRMKGENVATTEVAEVF 611
 DB 392 PQ-ATAKYLE-----ADGMLHTGDTGYRDEDFYFYVDRRCNMILKRGENVSCVELNII 445

QY 612 EALDFLOEVNVYGVTPGHEGRAGMALVLRPPHALDMLQLYTHVSENLPYARPEFLRL 671
 DB 446 AAHPKXQDIYVVGIR-K-DSTRDEAIKAFVYLINSEETLSEEPFRFCQNMAKRVPSTYSL 504
 QY 672 QESL 675
 DB 505 RKDL 508
 RESULT 9
 CALC_SALTI STANDARD; PRT; 517 AA.
 ID CALC_SALTI STANDARD; PRT; 517 AA.
 AC Q829J4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable crotonobetaine/carnitine-coA ligase (EC 6.3.2.-).
 GN CALC OR STY0081 OR T0072.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxId=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Barrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Tagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RA "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.P., Rose D.J.,
 RA Burdand V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RA "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -i- FUNCTION: Could catalyze the transfer of CoA to crotonobetaine
 CC or carnitine (By similarity).
 CC -i- PATHWAY: Carnitine metabolism (conversion of carnitine to
 CC gamma-butyrobetaine).
 CC CC
 CC -i- Similarity: Belongs to the ATP-dependent AMP-binding enzyme
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AL627265; CAD01225.1; -
 DR EMBL: AE016834; AA067805.1; -
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 KM Ligase; Complete proteome.
 SQ SEQUENCE 517 AA; 58509 MW; 547FC052A8FC1512 CRC64;
 Query Match 9.2%; Score 353.5; DB 1; Length 517;
 Best Local Similarity 24.4%; Pred. No. 2.6e-16;
 Matches 144; Conservative 84; Mismatches 248; Indels 113; Gaps 18;

```

QY 110 GGCSLAMRLAEIAQOAAHTFLIHGS-----RRFSYSEABERSNRAAPFLALGMDWGP 164
D 5 GGQNLQOMDDDLAAGVYGDXTALIFESCEGIVQFSYASINBEINFTALFY----- 55
QY 165 DGGDSGEGSAGEGERAPAGDAAAGSAGEFAGCGAARGGGAAPLSPGATVALLPAG 224
D 56 -----YLG-----IRKGRVALHLDNC 72
QY 225 PEFILMPLGAKAGLRTAFVPTALRRGPHLLHCLRSAGARVLAPEFL-----ESLEP 277
D 73 PEFICWFGLAKIGALMVPINARLLGSESAMLIQNSQVSLVTSQFPMREIRQDNST 132
QY 278 DDPALRANGHLMAAGPGRHPAGISDLAEVSAEVDGVPGLSSPGISITDTCLYIFTSG 337
D 133 PLNHICLIGEQLPADGVSH---FSQLQARQATL-----CYTPALSTDDTAELIFTSG 183
QY 338 TGLPKRAARISHLKLQCCGFYQL---CGVHQBVDVYLAFLYHMSGSLIGVCGMGAT 395
D 184 TTSRPGGVITHTN-LRFAGYSAMQIALRDDDVVTWVPFAHIDCCTAAMPASAGST 242
QY 396 VILKSFSAQGFWECCQHRVTVPQYIGELCRYLVNPPSKAERGHKVR-LAVSGLRPD 454
D 243 FVLEKYSARAFWDQVRKYQATVTECIPMTRTLWQVATPTDQHHLREVWFYLNLSAQ 302
QY 455 TWRFVRRRPGPLQVETYGLTGEGNATINTGQR-----GAVGRASMLYKHIFPFLI 507
D 303 EKDAFTERG-VRLITSYGMEITVGI---GDRPGDKRMPISIGRVGSYE----- 350
QY 508 RYDVTTGEBIRDPQGHCAVTSPEPG-LIVAPVSOQSPFLGAGPELAQKLDVFRP 566
D 351 -----AEIRDDQNRPLPA--GEIGELCIKIGIPKTIKEXYMOPE-ATAALAP----- 396
QY 567 GVPFNTGDLVCCDQGFIRFHDRTGDFRMKGENVATTEVAEVEALDFLOEVNYYGT 626
D 397 -EGWHLTGDSGYDEBGFYFVDRRCNMIKRGENVSCVELENIISAHPKIDIVVGIK 455
QY 627 VPGHEGRAGMALVLRPHALDLMQLYTVHSENLPYARPRFLRQESL 675
D 456 -DAIRDEAIKAFIVLNEGRTLSAEPFSFCENNAKFKVPSMEIRIDL 503

RESULT 10
CAIC_SALTY STANDARD; PRT; 517 AA.
AC Q8ZK4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Probable crotonobetaine/carnitine-CoA ligase (EC 6.3.2.-).
GN CAIC OR STM0071.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Jorvolik S., Ali J., Dante M., Du F., Hou S., Jayman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2";
RL Nature 413:852-856 (2001).
CC -1- FUNCTION: Could catalyze the transfer of CoA to crotonobetaine
CC or carnitine (By similarity).
CC -1- PATHWAY: Carnitine metabolism (conversion of carnitine to
CC gamma-butyrobetaine).
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.

```

```

CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008696; AAL1903.1; -.
DR StvGene; SG72727; caic.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PROSITE; PS00455; AMP-BINDING; 1.
KW Ligase; Complete proteome.
SQ SSQURCE 517 AA; 58487 MW; 7F980E9891C67634 CRC64;

Query Match 9.2%; Score 352.5; DB 1; Length 517;
Best local Similarity 24.6%; Pred. No. 3e-16;
Matches 145; Conservative 84; Mismatches 247; Indels 113; Gaps 18;

QY 110 GGCSLAMRLAEIAQOAAHTFLIHGS-----RRFSYSEABERSNRAAPFLALGMDWGP 164
D 5 GGQNLQOMDDDLAAGVYGDXTALIFESCEGIVQFSYASINBEINFTALFY-SLG----- 58
QY 165 DGGDSGEGSAGEGERAPAGDAAAGSAGEFAGCGAARGGGAAPLSPGATVALLPAG 224
D 59 -----IRKGRVALHLDNC 72
QY 225 PEFILMPLGAKAGLRTAFVPTALRRGPHLLHCLRSAGARVLAPEFL-----ESLEP 277
D 73 PEFICWFGLAKIGALMVPINARLLGSESAMLIQNSQVSLVTSQFPMREIRQDNST 132
QY 278 DDPALRANGHLMAAGPGRHPAGISDLAEVSAEVDGVPGLSSPGISITDTCLYIFTSG 337
D 133 PLNHICLIGEQLPADGVSH---FTQLQARQATL-----CYTPALSTDDTAELIFTSG 183
QY 338 TGLPKRAARISHLKLQCCGFYQL---CGVHQBVDVYLAFLYHMSGSLIGVCGMGAT 395
D 184 TTSRPGGVITHTN-LRFAGYSAMQIALRDDDVVTWVPFAHIDCCTAAMPASAGST 242
QY 396 VILKSFSAQGFWECCQHRVTVPQYIGELCRYLVNPPSKAERGHKVR-LAVSGLRPD 454
D 243 FVLEKYSARAFWDQVRKYQATVTECIPMTRTLWQVATPTDQHHLREVWFYLNLSAQ 302
QY 455 TWRFVRRRPGPLQVETYGLTGEGNATINTGQR-----GAVGRASMLYKHIFPFLI 507
D 303 EKDAFTERG-VRLITSYGMEITVGI---GDRPGDKRMPISIGRVGSYE----- 350
QY 508 RYDVTTGEBIRDPQGHCAVTSPEPG-LIVAPVSOQSPFLGAGPELAQKLDVFRP 566
D 351 -----AEIRDDQNRPLPA--GEIGELCIKIGIPKTIKEXYMOPE-ATAALAP----- 396
QY 567 GVPFNTGDLVCCDQGFIRFHDRTGDFRMKGENVATTEVAEVEALDFLOEVNYYGT 626
D 397 -EGWHLTGDSGYDEBGFYFVDRRCNMIKRGENVSCVELENIISAHPKIDIVVGIK 455
QY 627 VPGHEGRAGMALVLRPHALDLMQLYTVHSENLPYARPRFLRQESL 675
D 456 -DAIRDEAIKAFIVLNEGRTLSAEPFSFCENNAKFKVPSMEIRIDL 503

RESULT 11
LCFA_YERPE STANDARD; PRT; 562 AA.
AC Q8ZS9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3) (Long-chain acyl-CoA
DE synthetase).
GN FADD OR YPOC074 OR Y2236.
OS Yersinia pestis.

```


OC	Bacteria	Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC	Enterobacteriaceae; Yersinia	
OX	NCBI_TaxID=632;	
RN	(1)	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=CO-92 / Biovar Orientalis;	
RX	MEDLINE=21470413; PubMed=1158560;	
RA	Parkhill J., Wren B.W., Thomson N.R., Titchall R.C., Holden M.T.G.,	
RA	Prentice M.B., Sebaitia M., James K.D., Churcher C., Mangall A.L.,	
RA	Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Parraga A.M.,	
RA	Chillingworth T., Cronin A., Davies R.W., Davis P., Dougan G.,	
RA	Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,	
RA	Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,	
RA	Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.;	
RT	"Genome sequence of Yersinia pestis, the causative agent of plague.",	
RL	Nature 413:523-52(2001).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=KIMS / Biovar Mediaevalis;	
RX	MEDLINE=22137863; PubMed=12142430;	
RA	Degg W., Barland V., Plunkett G., III, Boulin A., Mayhew G.F., Iiss P.,	
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,	
RA	Petherston J.C., Lindler L.E., Brudaker R.R., Plano G.V.,	
RA	Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,	
RA	Perry R.D.;	
RT	"Genome sequence of Yersinia pestis KIM";	
RL	J. Bacteriol. 184:4601-4611(2002).	
CC	-I- FUNCTION: Esterification, concomitant with transport, of exogenous	
CC	long-chain fatty acids into metabolically active CoA thioesters	
CC	for subsequent degradation or incorporation into phospholipids (By	
CC	similarity).	
CC	-I- CATALYTIC ACTIVITY: ATP + a long-chain carboxylic acid + CoA = AMP	
CC	+ diphosphate + an acyl-CoA.	
CC	-I- COFACTOR: Magnesium (By similarity).	
CC	-I- SUBUNIT: Homodimer (Probable).	
CC	-I- SUBCELLULAR LOCATION: Partially membrane-associated (Potential).	
CC	-I- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme	
CC	family.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation	
CC	at the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/	
CC	or send an email to license@isb-sdb.ch).	
CC	-----	
DR	EMBL; AJ14151; AACG9886.1; -.	
DR	EMBL; AE013826; CAM65796.1; ALT_INIT.	
DR	PIR; AB0253; AB0253.	
DR	InterPro; IPR000873; AMP-bind.	
DR	Pfam; PF00501; AMP-binding_1.	
DR	PROSITE; PS00455; AMP_BINDING; 1.	
KW	Ligase; Fatty acid metabolism; Magnesium; ATP-binding; Membrane;	
KW	Complete proteome.	
NP_BIND	213	
FT	VARIANT 504 504	
FT	SEQUENCE 562 AA; 62639 MW; 812AB72B3J13D3FD CRC64;	
QY	Query Match	
Db	Best Local Similarity 7.2%; Score 276.5; DB 1; Length 562;	
Db	Matches 131; Conservative 69; Mismatches 234; Indels 95; Gaps 18	
QY	211 LSPGVATVLLPAGGEFLMTFGAKAGIRTAFTPTALRRGFLLHCLRSRGARALVLAPR 270	
Db	71 LQKGGRVALMPNNLLQYIFAIFGLVRAGMIIVANNPLYPRELREQLSDSGVAIVISN 130	
QY	271 FLESLEPDL-----PALRWGHLMAAGPGTHPAGISDLIAEVSAREVDGPVPGLSP 323	
Db	131 FAHTLEKVFKTVGNHVILTRNGDL-SAAKT-----LVNFVKYIKRLVPEKY-LP 181	
QY	324 QSI-----TDLCVLYFTSGTGLPKAARISH---LKINQQCGF 358	

Db	152	DAISRTVTLQGRKRNQYKRPDVINTDIAFLQYTGGTGVAKGAILTHRNQSNIEQAKYA	241
Qy	359	YQLCGVHGEDVLYIALPLVHNSGSLIGVCGMI-----GATVVLKSKFSAQGFMEDCOH	414
Db	242	YAPLLQPGRDVLVWTLPLVYHIFAL---TVNCLFLFELGREGSLITNPREDIPGVKLEISRY	298
Qy	415	RTVYQYVIGELCRVLYVNPSPKAKRGHVRVLAVSGSLR-----PDYWRFPARRGPELOVL	469
Db	299	PFTALTGNTVTFNNALNNEEFTHLDFTSLRUSVGGGMVQKAAVAKKETTNGK---HLL	354
Qy	470	ETVYGTGEGNVAITINYTGORGAVGRASMLYKHIFFPSILIRYDVTGSEPT-----RDPQG	522
Db	355	EGYGTETGESPVLV---TGN-----PYDLKHSYSGSLGLPSPSTDVRLRDDG	396
Qy	523	HOMATSPSEPGLLVAVPYSQSPFLGYAGPEPLAOGSKLKVDFRPQDVFFENTGDLVCDQD	582
Db	397	NDVEL--GKPEBLW--VRGPQVMGMYQRPDNT-----DVLKDG--WLAGTDIATWDED	445
Qy	583	GFLRPHDITGTFPMKGENVATTEVVAEFEALDPLQEVNIVYGVTPGHEGRAGMAALVLR	642
Db	446	GFLRVYDRKKMILVSGNPVNPNEIBEYVALHAAYLESAYIG--VPEHVSGEAYKVFVK	503
Qy	643	PPHALDMQLYTHVSENLPPYARPRFLRLQSLATTETFKQOXRVANE	691
Db	504	NDASITPEELLTHCRRLYLTGYKVAKIVEFRDELPSKNVKGILRRLEDE	552
RESULT 12			
ICFA_ECO57	1CFA_ECO57	STANDARD;	561 AA.
AD	Q8XDR6;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	long-chain-fatty-acid-CoA ligase (EC 6.2.1.3) (long-chain acyl-CoA synthetase).		
GN	PADD OR OLDID OR Z2848 OR ECS2514.		
OS	Escherichia coli O157:H7.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
OC	Enterobacteriaceae; Escherichia.		
OX	NCBI_TaxID=83334;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=O157:H7 / EDL933 / ATCC 700927;		
RC	MEDLINE=21074935; PubMed=11206551;		
RA	Perna N.T., Plunkett G., Irii, Burland V., Mau B., Glasner J.D.,		
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,		
RA	Grotbeck E.U., Davis N.W., Lim A., Dimalanta B.T., Potlunouis K.,		
RA	Apodaca V., Anantharaman T.S., Lin Y., Yen G., Schwartz D.C.,		
RA	Weich R.A., Blattner P.R.;		
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";		
RL	Nature 409:529-533(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=O157:H7 / RUMD 0509952;		
RC	MEDLINE=21156231; PubMed=11258796;		
RA	Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,		
RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,		
RA	Iida T., Takami H., Honda T., Sasaki K., Ogawara N., Yasunaga T.,		
RA	Kuhara S., Shiba T., Hattori M., Shinagawa H.		
RT	"Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12";		
RL	DNA Res. 8:11-22(2001).		
CC	-1- FUNCTION: Fatty acidification, concomitant with transport, of exogenous		
CC	long-chain fatty acids into metabolically active CoA thioesters		
CC	for subsequent degradation or incorporation into phospholipids (By		
CC	similarity).		
CC	-1- CATALYTIC ACTIVITY: ATP + a long-chain carboxylic acid + CoA = AMP		
CC	+ diphosphate + an acyl-CoA.		
CC	-1- COFACTOR: Magnesium (By similarity).		
CC	-1- SUBUNIT: Homodimer (Probable).		
CC	-1- SUBCELLULAR LOCATION: Partially membrane-associated (Potential).		

-1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE005403; AAG56794.1; -
 DR EMBL; AE005558; BAB35937.1; -
 DR PIR; B90943; B90943.
 DR PIR; F85791; F85791.
 DR InterPro; IPR000873; AMP-bind.
 DR Pfam; PF00455; AMP_BINDING.1.
 DR PROSITE; PS00455; AMP_BINDING.1.
 KW Ligase; Fatty acid metabolism; Magnesium; ATP-binding; Membrane;
 KW Complete proteome.
 FT NP BIND 213
 FT SEQUENCE 561 AA; 62364 MW; 4DB944AB7DF40CF2 CRC64;

Query Match 7.1%; Score 273.5; DB 1; Length 561;
 Best Local Similarity 24.2%; Pred No. 6.2e-11;
 Matches 125; Conservative 87; Mismatches 234; Indels 71; Gaps 18;

QY 211 LSPGATVALLPAGSEFLMFWGLAKAGRTAFVETALRQGLHCLSCGARRLVLAPE 270
 Db 71 LKGGPRVALMNMENLQYPAVLPGLRAGWIVVNPVLYPRELHQLNDSASAIIVISN 130
 QY 271 FLESLEP-DLPA-----LRAMGLHMAAG-----PGTHPGISDLAEV 308
 Db 131 FATHLEKVVDTAVOVHILTRMGDOLSTKGLVNVVVKYIKRLVPEKH---LPDAISFR 187
 QY 309 SAENVDPVPGYLSQSPITDTCLYI-FTSGTGLPKAARISHKTL-----OCQGYQCG 363
 Db 188 SALHNGYRQGY-KPELVPEDLAFQYGTGTGVAKGAMLTFRNMLAMEGVNATYGPDL 246
 QY 364 VHOEDVITYALPLVHMSGSLGIVGCMGGA-TVVLKSKFSAGQWECOCQHRVTPGYI 422
 Db 247 HPGKEIVLVAFLPLVHIFALTINCLFELGGONLITFRDIPGLVKEKLAKYFPAITGV 306
 QY 423 GELCRVLYNPSPSKERGHKVLAVAGSL--RPDTERFVRFRFGQLVLETYGLTE-GNV 479
 Db 307 NTLFNALLNKKERQQLDESSLHSAGCGPVOQVAERVKLTGQ-YLLEGGLTECAPL 365
 QY 480 AITNTYGGAGAVGRASWLVKHLFPSLRIVDTTSEPIRDPQGHCM-----ATSPGEFGL 534
 Db 366 VSVN-----PYDIDYHSGSIGLPVPSSTAKLVDDDNDEVSPGQGE 406
 QY 535 IYAPVSGQSPFLGYAGPELAGSKLLKDVFRGADVFPMTGDLVDDQGFRLFRHRTGDT 594
 Db 407 LC-VAGPQVLMGYWQRPD-ATDEITK-----NGMLHTGDLAVWDEBGLRIYDRKDM 457
 QY 595 FRKKGENVATTEVAEFAELDLQEVNVVYVTPVGHGEGAGAAVLVLRPHALDLMQLYT 654
 Db 458 ILVSGNVVYVNEIEDVVMQHPGVQVAAVGVP-SSSGBAVKI FVVKDPS-SLTBSISVT 515
 QY 655 HUSENPPYARPRFLQESLATTETFRQKRYRMANE 691
 Db 516 FCRRLQLGKVPKLVFREDLPKSNVYKTLRRELDE 552

RESULT 13
 LCRA_ECOLI STANDARD; PRT; 561 AA.
 AC F29212;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Long-chain-fatty-acid-CoA ligase (EC 6.2.1.3) (Long-chain acyl-CoA synthetase).

GN FADD OR OLD OR B1805 OR C2209.
 OS Escherichia coli, and
 OS Escherichia coli 06.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 CX NCBI_TaxID=562, 217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=94150456; PubMed=8107670;
 RA Fulda M., Heinz E., Wolter F.P.;
 RT "The fadD gene of Escherichia coli K12 is located close to rnd A
 RT 39.6 min of the chromosomal map and is a new member of the
 RT AMP-binding protein family";
 RL Mol. Genet. 242:241-249(1994).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
 RC STRAIN=K12;
 RX MEDLINE=93094273; PubMed=1460045;
 RA Black P.N., DiRusso C.C., Metzger A.K., Helmert T.L.;
 RT "Cloning, sequencing, and expression of the fadD gene of Escherichia
 RT coli encoding acyl coenzyme A synthetase";
 RL J. Biol. Chem. 267:25513-25520(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden W.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1233-1238(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251358; PubMed=9097040;
 RA Itoh T., Alba T., Fujita K., Hayashi K., Inada T.,
 RA Iono K., Kasai H., Kimura S., Kikagawa M., Kikagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
 RA Sato N., Sempel G., Seki Y., Sivaraman S., Tagami H.,
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horichi T.;
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 40.1-50.0 min region on the linkage map";
 RL DNA Res. 3:379-392(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=06:HI / CFT073 / ATCC 700928;
 RX MEDLINE=2388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,
 RA Raeko D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blatner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic Escherichia coli";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 RN [6]
 RP IDENTIFICATION OF ATP-BINDING MOTIF, AND MUTAGENESIS OF TYR-213;
 RX THR-214; GLY-216; THR-217; GLY-219; LYS-222 AND GLU-361.
 RP MEDLINE=22162442; PubMed=12034706;
 RA Weimar J.D., DiRusso C.C., Delio R., Black P.N.;
 RT "Functional role of fatty acyl-coenzyme A synthetase in the
 RT transmembrane movement and activation of exogenous long-chain fatty
 RT acids. Amino acid residues within the ATP/AMP signature motif of
 RT Escherichia coli fadD are required for enzyme activity and fatty acid
 RT transport";
 RL J. Biol. Chem. 277:29369-29376(2002).
 CC -1- FUNCTION: Esterification, concomitant with transport, of exogenous
 CC long-chain fatty acids into metabolically active CoA thioesters
 CC for subsequent degradation or incorporation into phospholipids.
 CC -1- CATALYTIC ACTIVITY: ATP + a long-chain carboxylic acid + CoA = AMP
 CC + diphosphate + an acyl-CoA.

```

CC -1- COFACTOR: Magnesium.
CC -1- SUBUNIT: Homodimer (Probable).
CC -1- SUBCELLULAR LOCATION: Partially membrane-associated (Potential).
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme family.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL, X70994; CA50321.1; -
DR EMBL, I02649; AAA23752.1; -
DR EMBL, AE000275; AAC74875.1; -
DR EMBL, D90823; BAA15600.1; -
DR EMBL, D90824; BAA15609.1; -
DR EMBL, D90825; BAA15614.1; -
DR EMBL, AE016761; AAN80668.1; ALT_INIT.
DR PIR, E64941; S41589.
DR HSSP, P08659; ILCT.
DR Ecogen, EG11530; fadd.
DR InterPro, IPR000873; AMP-bind.
DR Pfam, PF00501; AMP-binding; 1.
DR PROSITE, PS00455; AMP BINDING; 1.
KW Ligase; Fatty acid metabolism; Magnesium; ATP-binding; Membrane; Complete proteome.
FT NP_BIND 213 224 ATP (PROBABLE).
FT MUTAGEN 213 213 Y->A: LOSS OF ACTIVITY.
FT MUTAGEN 214 214 T->A: 10% OF WILD-TYPE ACTIVITY.
FT MUTAGEN 216 216 G->A: DECREASE IN ACTIVITY.
FT MUTAGEN 217 217 T->A: DECREASE IN ACTIVITY.
FT MUTAGEN 219 219 G->A: DECREASE IN ACTIVITY.
FT MUTAGEN 222 222 E->A: DECREASE IN ACTIVITY.
FT MUTAGEN 361 361 E->A: LOSS OF ACTIVITY.
FT CONFLICT 34 51 ARYADQPAFVNMGEVYTF -> GALRSTWCVEYGGNDL (IN REF. 2).
FT CONFLICT 468 490 NEIDVYMQHGVQDEVAAYGPS -> TRLKMSQSMVAYRKSRLALVP (IN REF. 2).
FT CONFLICT 496 496 A -> G (IN REF. 2).
FT CONFLICT 555 555 GVNDYKA -> QSGQ (IN REF. 2).
SQ SEQUENCE 561 AA; 62332 MW; 249B0AA54B3DEFA5 CRC64.

Query Match
Best Local Similarity 24.0%; Pred. No. 1,2e-10;
Matches 124; Conservative 87; Mismatches 235; Indels 71; Gaps 18;

7.0%; Score 269.5; DB 1; Length 561;
Query 211 LSPGATVALILDPAGEFLIMWFGGLAKGLRTAFVPTALRQGLHLCHRCGARRLVLAPE 270
Db 71 LKKGKRVALLMMENLQYPAVLFGLIRAGMIVVNVPLTYPRELHQNDLSASAIIVYSN 130
271 FLESDEP--DIPA-----LRAGGLHMAAG-----PGTHPAGISDLAEV 308
Db 131 FAHTEKVVYDKRAVGHVILTRMGDQLSTAKGTIVNVFVVKYIKRLVFKTH--LIDALISFR 187
309 SAEVDGVPVPGYISSPQSTITDTCLYI-FTSGTTGLPKAARISHLKL----QCQGFYQJG 363
Db 188 SALHNGHYMOQYV-KPELVPEIDALFQYGGTGVAKGMLTTHRMNLALIEQVNAITYGLL 246
364 VHQEIVITLALPLVYMGSSILGIVGCMGIGA-TVVLKSKFSAQGWEDCCQGHRTVPOYI 422
Db 247 HPGKLVVVTALPLVYHITFALTINCLFTEHGQNTLITTPRDIPLGLVSKLAKYPTATIGV 306
423 GELCRVLYVNPSPSKAERGHKVLAVGSL--RPDTMERFVRFGPLQVLYETGYLLE-GNV 479
Db 307 NTLFNAALNNKRFQQLDFSSHLISAGGMPQGVQVAERWVLTQG-YLLEBYGGLTECAPL 365
480 ATINTTQGRGAVGRASMLYKHIFFPSLRIVDTTGEPIFRDQGHM-----ATSPGEFGL 534
Db 366 VSVN-----PYDIDYHSGSIGLPVPSTEARKLAVDDDNNEVPPQGPGE 406

```

```

QY 535 LVAPVQSPFPLGAYGPELTAQGLLKDVPRPDVFNFTGDLIVCDQGFIRHEDTGP 594
Db 407 LC--VKGPOVWLKYWOPD-ATDEIIR-----NGWLHTGDIADWBEGLRIVDKKDV 457
QY 595 FRWKGENVATTEVAEVEFALDFQEVNVVGVTVPGHEGAGMAALVLRPHALDLMLQVLT 654
Db 458 ILVSGFVYVNEIEDVVMQHGPQVEVAAGVP-SGSSGEAVKLFVVKOP-SLTESLVLT 515
QY 655 HUSENLPVYRPPRFLQESLATTERPKQCKVMANE 691
Db 516 FCRRLQGVYKVPRLVBERDELPRSNVKGILRRLRDE 552

RESULT 14
LCFA_SALTY
ID LCFA_SALTY STANDARD; PRT; 561 AA.
AC Q8XG38;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE long-chain-fatty-acid--CoA ligase (EC 6.2.1.3) (Long-chain acyl-CoA synthetase).
GN FADD OR STM1818 OR STM1948 OR T1059.
OS Salmonella typhimurium, and
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxId=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SSGC1412 / ATCC 700720;
RX MEDLINE=21534947; Pubmed=11677609;
RA Mclelland M., Sanderson K.E., Spieth J., Clifton S.W., Lettelle P., Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D., Leonard S., Nguyen C., Scott K., Holmes A., Greal N., Mulvaney E., Ryan E., Sun H., Floria L., Miller W., Stoneking T., Nhan M., Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2";
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=CT18;
RX MEDLINE=21534947; Pubmed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M., Baker S., Basham D., Brooks K., Chillingworth T., Connor P., Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J., Feldwell T., Hamlin N., Hogue A., Hien T.T., Holroyd S., Jagsis K., Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C., Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RA "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium CT18";
RL Nature 413:848-852(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; Pubmed=12644504;
RA Deng W., Liou S.-R., Plunkett G.II, Mayhew G.F., Rose D.J., Burland V., Kodoyianni V., Schwartz D.C., Blatter F.R.;
RA "Comparative genomics of Salmonella enterica serovar Typhimurium strains Ty2 and CT18";
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- FUNCTION: Esterification, concomitant with transport, of exogenous long-chain fatty acids into metabolically active CoA thioesters for subsequent degradation or incorporation into phospholipids (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + a long-chain carboxylic acid + CoA = AMP + diphosphate + an acyl-CoA.
CC -1- COFACTOR: Magnesium (By similarity).
CC -1- SUBUNIT: Homodimer (Probable).
CC -1- SUBCELLULAR LOCATION: Partially membrane-associated (Potential).

```

CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: AE008781; AL20733.1; -
 CC EMBL: AL627272; CAD05501.1; -
 CC EMBL: AE016837; AA068725.1; -
 CC Styrene; SGG7777; faad.
 CC Interpico; IP000873; AMP-bind.
 CC Pfam; PF00501; AMP-binding; 1.
 CC PROSITE; PS00455; AMP BINDING; 1.
 CC Lysase; Fatty acid metabolism; Magnesium; ATP-binding; Membrane;
 CC Complete proteome.
 CC NP BIND 213
 CC SEQUENCE 561 AA; 62223 MW; 4000B2948FE3D15 CRC64;

Query Match 6.9%; Score 265.5; DB 1; Length 561;
 Best Local Similarity 24.0%; Pred. No. 2.1e-10;
 Matches 124; Conservative 80; Mismatches 243; Indels 69; Gaps 15;

QY 211 LSPGATVALLPAGPEFLMWFGKAGRTAFVPTALRGGLHCLSCGABLLVLAPE 270
 DB 71 LKKGRVALLMPNLLQYVPAFGIIRAGIVVNVPLTPRELEQLDMSGAALITNSN 130
 QY 271 PLESLEPDLP-----ALRANGHLHMAAG-----PCTHPAGISDILA 308
 DB 131 PAHLEKVEKTSVGHVILTRMGDDLTAKGTAVNVFVKYIKRLVEXKH---LPDAISFR 187
 QY 309 SAEDVGPVGYSSPSQSDITDTCLYFTSTTGLPKAAISHLKLT---CCQGFYQLCGV 364
 DB 188 SALOHGYRMQYKPEVVAEDLAFLOYTGTGVAAGALTRHNMALNEQVATYGPILH 247
 QY 365 HOEDVYLLPLPYHMSGLIGVCGMGIGA-TVVLKSKFSAQGFEDCCQHRVTFQYIG 423
 DB 248 PKELELVTLPLPYHIFALTMNGLLPIELGGNLLITNRPDIPLGLKELAKPFTAMGVN 307
 QY 424 ELCRYLVNPPSKAERGHKVLAVSGSL--RPDTWERFYRRRGPLOVETGLTE-GNVA 480
 DB 308 TLFNALLNKKRQQLDFSSILHSAGGMPQVQVNAERWKLGTQ--YLLEGVGLTECAPLV 366
 QY 481 TINYTGQRGAVRASMLYGHIFPFSLRVDVTTGPIRDPQCHM-----ATSPGPEGLL 535
 DB 367 SVN-----PHDIDHSGSIGLPVSTKAKLVDDDDNVAVGEAGEL 407
 QY 536 VAPVSGQSPFLGYAGPELAQGLKLDVFRPGDVFENFGDILLVCDQGFLEPHRTGDTF 595
 DB 408 C--VKSPQVLMGWGRPD-ATDEIRK-----DGMHTGDIAVMDEDFLRIVDRKKMI 458
 QY 596 RKGKENVATTEVAEVEALDFLOENVVGVTVPGHEGRAGMAALVLRPHALDLMQVLT 655
 DB 459 LVSGFENVYNELEDVMOHSGVQVEVAAGVP--SGSSGEVVKLFVVKDP--ALTDALITF 516
 QY 656 VSENLPPYARPRFLRLOESLATTETFEKQKVMANE 691
 DB 517 CRRLHLYGVPRKQVRELFELPKSNVGLKLRRLRDE 552

RESULT 15
 MENE ENTRE STANDARD; PRT; 485 AA.
 AC 0838K1;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE O-succinylbenzoate-CoA ligase (EC 6.2.1.26) (OSB-CoA synthetase) (O-succinylbenzoate-CoA synthetase).

GN MENE OR EF0446.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NOBI_TaxID=1351;

RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RA MEDLINE=22550857; PubMed=12663927;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Kettelin H., Dodson R.J., Umayam L., Brinkac L., Beaud M.,
 RA Dughery S., DeBoy R.T., Durkin S., Kolony J., Madupu R., Nelson W.,
 RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
 RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
 RA "Role of mobile DNA in the evolution of vancomycin-resistant
 RT Enterococcus faecalis";

RL Science 299:2071-2074(2003).
 CC -1- FUNCTION: Converts O-succinylbenzoate (OSB) to O-succinylbenzoate + CoA = AMP + CoA (OSB-CoA).
 CC -1- CATALYTIC ACTIVITY: ATP + O-succinylbenzoate + CoA = AMP +
 CC -1- dihydrophosphate + O-succinylbenzoate + CoA.
 CC -1- PATHWAY: Menadiquinone biosynthesis.
 CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme family. Mene subfamily 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: AE016948; AA080301.1; -
 CC TIGR: EF0446; -
 DR HAMAP: MF 00731; -; 1.
 DR InterPro: IPR000873; AMP-bind.
 DR Pfam: PF00501; AMP-binding; 1.
 DR PRINTS: PR00154; AMPBINDING.
 DR TIGRFAMs: TIGR01923; mene; 1.
 DR PROSITE: PS00455; AMP BINDING; 1.
 KM Menadiquinone biosynthesis; Lysase; Complete proteome.
 SQ SEQUENCE 485 AA; 54289 MW; 1BAD82BC8C0882 CRC64;

Query Match 6.7%; Score 259; DB 1; Length 485;
 Best Local Similarity 23.8%; Pred. No. 4.9e-10;
 Matches 119; Conservative 71; Mismatches 213; Indels 96; Gaps 18;

QY 217 VALLPAGPEFLMWFGKAGRTAFVPTALRGGLHCLSCGABLLVLAPEPLESLE 276
 DB 51 VALFSKSKELVYSILALWELGKELLFLNTHLTLAELTQLDAQVKTIGAPE----- 104
 QY 277 PDLPALRANGHLHMAAGPCTHPAGISDILA 308
 DB 105 -----TQALLLEISFVDVQCMIKKSHLSHOEFQSPSDLS 140
 QY 329 TCIYFTSGTGLPKAARISHLK--LLCCQGFYQLCGVQEDVYLLALPLPYHMSGLIGI 386
 DB 141 VASIMTSGTGPQKAV--LQRFKNHLASARGIOENMGITABDCMLCAVLFHISG--LSLV 198
 QY 387 VGCWGIGATVLLKSKPSAQGFEDCCQHRVTVFQYIGELCRYLVNPPSKAERGH--K 442
 DB 199 VRQLTVAGCSIRLYDKFDEQVOTDLQEGRGTVISVVAITMLQQLLSVYP--EAGYASPK 255
 QY 443 VRLAVSGRLRPDWEFRFRFGLOVLTETGLTE--GNVATINYQQRGAVRASMLYH 500
 DB 256 GMLDGGPIAPDLACCEKGIIP--VIGSYGMEITCSQVVALKFEEDALKITGSA----- 307
 QY 501 IFFPSLRVDVTTGPIRDPQCHWATSPEPEGLLAVPVSGQSP--FLG--YAGPELAQ 556
 DB 308 -----GQLKXMQIK-----IYDELGGQGPPEKQVGEIILKGNVVS 343
 QY 557 GKLLKQVFRP---GVFVNTGDLVLCDDQGFLEPHRTGDTFRKGENVATTEVAEVE 612

Db 344 GYL--NORQPEKWTADGWFKTGDMGYLDAQSYLYVSRLSLSELISGGENITYPTEVEQYLQ 401
Qy 613 ALDPLQEVNXYGVTVFGHEGRAGMALVLRPPHALDLMQLYTHVSENI.PPYARPRFLRLQ 672
Db 402 AITGIXKAAVVG--EPDAQMGAVPVAYVI-SDQETTLAQIODCCSRKLAKYKRPKRIYFC 458
Qy 673 ESLATTEFFKQOKVMANE 691
Db 459 HSFPQTASGKTKAKRPMTE 477

Search completed: June 30, 2004, 17:33:53
Job time : 19 secs

This Page Blank (uspto)